ETF DIVERGANCE FROM NET ASSET VALUE

Michael Mancini
Spring 2016

A thesis submitted in partial fulfillment of the requirements for baccalaureate degrees in Finance & Economics with honors in Finance

Reviewed and approved* by the following:

David Haushalter
Professor of Finance
Thesis Supervisor

Brian Davis
Professor of Finance
Honors Adviser

* Signatures are on file in the Schreyer Honors College.
ABSTRACT

One of the fundamental financial theories is the law of one price, or the idea that an asset will sell at the same price in all markets. Closed end funds (CEFs) often times appear to violate this law by trading at a different price than their stated net asset value (NAV). Those funds, however, have barriers that prevent arbitrageurs from participating and eliminating inefficiencies.

Exchange traded funds (ETFs) are structured differently than closed end funds in such a way that makes arbitrage possible. In times of financial stress, however, ETFs also can trade away from their net asset value. This paper explores the causes of ETF divergence from NAV and attempts to provide quantitative predictions as to when it occurs.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ iii

LIST OF TABLES ........................................................................................................... iv

ACKNOWLEDGEMENTS ............................................................................................... v

Chapter 1 Closed End Funds ....................................................................................... 1
  History of Closed End Funds ..................................................................................... 2

Chapter 2 ETF Overview .............................................................................................. 4
  History of Exchange Traded Funds .......................................................................... 6

Chapter 3 Prior Research on CEF Divergence ............................................................. 7
  Rational Explanations For CEF Divergence From NAV ........................................ 8
  Irrational Explanations For CEF Divergence From NAV ...................................... 11

Chapter 4 Prior Research on ETF Divergence ............................................................ 15
  Rational Explanations For ETF Divergence From NAV ........................................ 16
  Irrational Explanations For ETF Divergence From NAV ...................................... 17
  Other Research on ETF Divergence ....................................................................... 19

Chapter 5 Process ....................................................................................................... 22
  Fund Selection ......................................................................................................... 22
  Specifications ......................................................................................................... 25
  Generated Fields .................................................................................................... 26
  Measuring Premiums / Discounts and Deviations ................................................... 28
  Developing a Regression to explain Premiums/Discounts and Deviations from NAV . 32
  Defining the Explanatory Variables ..................................................................... 33

Chapter 6 Establishing that Discounts and Premiums Exist ......................................... 38

Chapter 7 Single Variable Explanations .................................................................... 45
  Explaining Premiums and Discounts on ETFs ......................................................... 46
  Explaining Deviations on ETFs ............................................................................. 56

Chapter 8 Multivariable Explanations ....................................................................... 66
Explaining Premiums and Discounts on ETFs .......................................................... 66
Explaining Deviations on ETFs ............................................................................. 71

Chapter 9 Conclusions ......................................................................................... 76
Appendix A Other Explanatory Variables ............................................................. 79
Appendix B Other Single Variable Explanations for Real Premiums / Discounts .. 89
Appendix C Other Single Variable Explanations for Deviations ......................... 137
BIBLIOGRAPHY .................................................................................................... 189
LIST OF FIGURES

Figure 1. Creation-Redemption System ................................................................. 5
Figure 2 The BDL Sentiment Index ................................................................. 13
Figure 3. Defining Real Premiums / Discounts ............................................... 30
Figure 4. Illustrating Executable Arbitrage in ETF Market ................................. 31
Figure 5. Average Closing Premiums / Discounts ............................................ 39
Figure 6. Average Bid-Ask Mid Premiums / Discounts ...................................... 40
Figure 7. Average Real Premiums / Discounts .................................................. 41
Figure 8. Average Closing Deviations ............................................................... 42
Figure 9. Average Bid-Ask Mid Deviations ....................................................... 43
Figure 10. Average Real Deviations ................................................................. 44
Figure 11. Average Financials Cost of Equity vs Average Premiums / Discounts Regression Output ................................................................. 47
Figure 12. Average Financials Cost of Equity vs Average Premiums / Discounts Model Output ................................................................. 47
Figure 13. Average Financials Cost of Equity vs Average Premiums / Discounts Model Output Error ................................................................. 48
Figure 14. Average Financials Cost of Equity vs Average Premiums / Discounts Scatterplot 48
Figure 15. Average Dividend Yield vs Average Premiums / Discounts Regression Output ... 49
Figure 16. Average Dividend Yield vs Average Premiums / Discounts Model Output ........ 50
Figure 17. Average Dividend Yield vs Average Premiums / Discounts Model Output Error 50
Figure 18. Average Dividend Yield vs Average Premiums / Discounts Scatterplot ........ 51
Figure 19. ETF Average Volume vs Average Premiums / Discounts Regression Output ..... 52
Figure 20. ETF Average Volume vs Average Premiums / Discounts Model Output ........ 52
Figure 21. ETF Average Volume vs Average Premiums / Discounts Model Output Error .... 52
Figure 22. ETF Average Volume vs Average Premiums / Discounts Scatterplot ........ 53
Figure 23. S&P Change vs Average Premiums / Discounts Regression Output .......... 54
Figure 24. S&P Change vs Average Premiums / Discounts Model Output ....................... 54
Figure 25. S&P Change vs Average Premiums / Discounts Model Output Error ................ 55
Figure 26. S&P Change vs Average Premiums / Discounts Scatterplot .......................... 55
Figure 27. Average Financials Cost of Equity vs Average Deviations Regression Output .... 56
Figure 28. Average Financials Cost of Equity vs Average Deviations Model Output ........ 57
Figure 29. Average Financials Cost of Equity vs Average Deviations Model Output Error .. 57
Figure 30. Average Financials Cost of Equity vs Average Deviations Scatterplot .......... 58
Figure 31. Average Dividend Yield vs Average Deviations Regression Output ................... 59
Figure 32. Average Dividend Yield vs Average Deviations Model Output ...................... 59
Figure 33. Average Dividend Yield vs Average Deviations Model Output Error ............. 60
Figure 34. Average Dividend Yield vs Average Deviations Scatterplot .......................... 60
Figure 35. ETF Average Volume vs Average Deviations Regression Output ..................... 61
Figure 36. ETF Average Volume vs Average Deviations Model Output .......................... 61
Figure 37. ETF Average Volume vs Average Deviations Model Output Error ................... 61
Figure 38. ETF Average Volume vs Average Deviations Scatterplot .............................. 62
Figure 39. Absolute Value of S&P Change vs Average Deviations Regression Output ....... 63
Figure 40. Absolute Value of S&P Change vs Average Deviations Model Output ............. 64
Figure 41. Absolute Value of S&P Change vs Average Deviations Model Output Error ...... 64
Figure 42. Absolute Value of S&P Change vs Average Deviations Scatterplot ................... 65
Figure 43. Theoretical Explanations for Premiums / Discounts With Prior Day Premium / Discount Regression Output ................................................................. 66
Figure 44. Theoretical Explanations for Premiums / Discounts With Prior Day Premium / Discount Model Output .................................................................................. 67
Figure 45. Theoretical Explanations for Premiums / Discounts With Prior Day Premium / Discount Model Output Error ................................................................. 67
Figure 46. Theoretical Explanations for Premiums / Discounts Without Prior Day Premium / Discount Regression Output ................................................................. 69
Figure 47. Theoretical Explanations for Premiums / Discounts Without Prior Day Premium / Discount Model Output
........................................................................................................................................ 69

Figure 48. Theoretical Explanations for Premiums / Discounts Without Prior Day Premium / Discount Model Output Error
............................................................................................................................................... 70

Figure 49. Theoretical Explanations for Deviations With Prior Day Deviation Regression Output
.................................................................................................................................................. 71

Figure 50. Theoretical Explanations for Deviations with Prior Day Deviation Model Output 71

Figure 51. Theoretical Explanations for Deviations with Prior Day Deviation Model Output Error
.................................................................................................................................................. 72

Figure 52. Theoretical Explanations for Deviations without Prior Day Deviation Regression Output ................................................................. 74

Figure 53. Theoretical Explanations for Deviations without Prior Day Deviation Model Output ................................................................................................. 74

Figure 54. Theoretical Explanations for Deviations without Prior Day Deviation Model Output Error........................................................................................................ 75
LIST OF TABLES

Table 1. ETFs Under Examination................................................................. 38

Table 2. Summary of Single Variable Explanations For Premiums and Discounts .......... 76

Table 3. Summary of Multivariable Explanations For Premiums and Discounts ............... 76

Table 4. Summary of Single Variable Explanations For Deviations .......................... 77

Table 5. Summary of Multivariable Explanations For Deviations ............................ 77
ACKNOWLEDGEMENTS

I would like to thank Professor Haushalter for guiding me through the thesis process and introducing me to new perspectives on fund divergence from net asset value.
Chapter 1

Closed End Funds

A closed end fund (CEF) is a publicly traded fund that holds a set amount of assets. The fund raises initial capital through an IPO, it usually does not issue follow-on shares, and funds do not buy back shares unless prearranged. Closed end funds do not redeem shares at the end of the day as an open-end fund does, rather investors must trade the ownership of the CEF to another interested party.

The Closed End Fund Puzzle is the study of the CEF’s divergence from net asset value (NAV). That is to say, that a fund holding $100 dollars of assets is not always valued on the market at $100 dollars. Instead, the value of the fund diverges and can trade at either a discount (below its net asset value) or a premium (above its net asset value). A good deal of research has covered reasons why investors may choose to over or under pay for the funds, and one reoccurring theme is that investors, especially retail investors, can be irrational. The CEFs have no structural feature that allows sophisticated investors to “correct” the price, so discounts and premiums can persist. Without informed investors able to participate, CEFs discounts and premiums could remain for extended periods of time.

Shortcomings in the CEF market led to the development of alternatives that could track the value of the underlining securities better. Open-ended funds were the first response. They allowed investors to redeem shares at value with NAV at end-of-day, but they did not allow...
investors to trade the fund interday. Exchange Traded Funds (ETFs) were the next innovation that appeared to give investors the trading flexibility of a CEF with the tracking ability of an open-end fund. An ETF is similar in creation and exchange to a CEF, with one major difference. An ETF can issue and buy back shares (in exchange for the portfolio of underlying securities) using the creation-redemption system. This feature allowed institutional investors to step in and stop funds from trading higher or lower than their NAV.

Despite the ETF’s creation-redemption system, there are times when the funds can trade below or above their NAV in the same way that CEFs do. These deviations are unique because, under the creation-redemption framework, institutional investors should have a nearly pure arbitrage opportunity, but they do not execute on it to the point of complete elimination of the inefficiency. Research on ETF divergence is sparse, but because of the similarities to CEFs, there is reason to believe that the research on Closed End Fund Puzzle can offer insights into this new ETF Enigma.

**History of Closed End Funds**

Closed end funds originated from the English investment trusts in the 1860s (Gabelli, 2004). The trusts primary investment targets were rail and construction ventures in the United States (Gabelli, 2004). Upper class individuals were the investors in these primitive CEFs (Hutson 2003). The market cap of these English investment trusts grew slowly and by 1903 it reached 70M pounds (Hutson 2003). The earliest trusts differed from CEFs today because
portfolio managers were forbidden from buying or selling securities without explicit notification to shareholders (Hutson 2003).

By the 1920s, the CEFs became popular with the American public with their total value estimated at $4.5B (Gabelli, 2004). Pre Depression, these funds traded at substantial premiums. In 1929, for example, the average premium on a closed end fund was 47% (Lee, 1990). In fact, prior to 1929, closed end funds almost always traded at a premium due to the public’s belief that the fund was worth the total value of assets plus the “manager’s skill” (Lee, 1990). It is important to note the closed end funds were not required to regularly post their holdings, so data from this time period can be misleading (De Long, 1990).

Following the 1929 crash, closed end funds underwent greater regulatory scrutiny due to Securities Act 1933 & 1934 and the Investment Company Act of 1940 (Gabelli, 2004). These regulations limited the amount of leverage the funds were allowed to take on and gave investors more transparency on the contents. Postwar, perception of CEFs changed. Instead of trading at a premium all the time, funds fluctuated from a 5% premium to a 25% discount relative to NAV (De Long, 1990).

For the 1965-1985 period, the closed end funds traded at an average discount of 10%, but the discounts themselves varied wildly (Lee, 1990). Research from the same time frame indicated that the discounts on the largest fund, Tricon, fluctuated between a 2.5% premium and a 25% discount (Lee, 1990). The discounts from Tricon were correlated with the discounts of the largest eight other funds with a coefficient of .5 (Lee, 1990). This research indicates that CEF
discounts and premiums move in tandem. The discounts followed a roughly seasonal pattern and were correlated monthly with a coefficient of .2-.4 (Lee, 1990).

In the late 1980s, specialized closed end funds became an emerging investment tool (Gabelli, 2004). These new closed end funds focused less on the broad market, and more on specific regions (Gabelli, 2004). International closed end mutual were designed to give investors exposure to more unique asset classes. Although, these new funds offered very specialized investment opportunities to investors, general interest among the public began to decline. Competition from open-end funds and exchange traded funds largely curtailed growth in the closed end fund market in the late 1990s to the present.

Chapter 2
ETF Overview

An exchange traded fund (ETF) is a publicly traded fund that can hold a varying amount of assets. ETFs are created by an institutional manager, or a fund sponsor, who aggregates the funds initial “creation units” or blocks of securities the ETF holds. Unlike a CEF, an ETF can issue new shares to special “authorized participants” in exchange for securities. Additionally, an ETF can “buy back shares” by paying “authorized participants” a set amount of securities that reflects the ETF’s portfolio. All “authorized participants” are financial institutions who have entered a legal contract with the fund’s manager to create and redeem ETF shares. The process of indirectly buying back and selling ETF shares is the creation-redemption system.
When an ETF is trading for more than the value of its underlining securities, authorized participants send the ETF a portfolio of those securities and the ETF would issue new shares to the authorized participant. Since the portfolio of the securities was valued less than the ETF, the authorized participant profits from the spread. The same situation holds true in the redemption process. If an ETF is valued less than the portfolio of securities that it holds, the authorized participants send back ETF shares in exchange for the portfolio and profit the difference in value. So, whenever an ETF is trading above or below its net asset value, authorized participants should have a pure arbitrage opportunity, assuming they are able to replicate the ETF portfolio instantaneously.
**History of Exchange Traded Funds**

ETFs were first created as a concept by the United States government in response to the Black Monday crisis in 1988 (Aoki, 2016). The idea behind the ETF was to provide a “liquidity buffer” to protect the underlining markets from unnecessary trading while allowing investors to still get exposure to assets (Aoki, 2016). Since CEFs had no mechanism to correct for mispricing, the ETF, along with the creation-redemption system, were developed to track the underlining portfolio with fewer errors. ETFs were first used primarily by institutional investors to hedge index positions and to keep cash active (Dow, 2006). They quickly gained popularity with the retail market and can account for up to 30% of all retail exposure today (Dow, 2006). The market currently stands at $3T dollars and is still growing (Aoki, 2016).

One of the primary reasons for this growth is due to the fact that ETFs track major indices with significant less NAV deviations than their CEF counterparts. The premiums and discounts on ETFs are usually less than 1%. This is because ETFs actively issue and buy back shares whenever the ETF diverges too much from the NAV using the creation-redemption system.
Chapter 3

Prior Research on CEF Divergence

The defining feature of a closed end fund is its tendency for its market value to diverge from the stated net asset value of the fund’s components. Closed end funds represent an apparent contradiction to the efficient market hypothesis (Lee, 1990). The market should value the closed end fund and its components equally or else, rational investors would immediately transfer holdings from one market to the other until the closed end fund and the retail markets equilibrate (Lee, 1990). Equilibration, however, usually only happens when a CEF is closed down and all of its assets are sold. When a fund is terminated (or its termination is announced), it tends to converge to its NAV, which should be expected (Lee, 1990). Managers of the funds are hesitant to terminate and often include anti-dissolution language in the fund’s charter (Lee, 1990).

In some studies, this discrepancy between the NAV and current valuation of closed end funds enables investors to achieve above-average returns (Thompson, 1978). Using data from closed end funds from 1940 to 1975, research indicated that purchasing discounted funds enabled investors to achieve a higher risk-adjusted return as opposed to holding a benchmark (Thompson, 1978).

More recent research from the 1965-1985 period verifies prior assumptions that buying closed end funds beat risk-adjusted benchmarks (Pontiff, 1995). This analysis moves one step further by adding transactional costs into the equation and provides a statistically significant strategy that investors can use to gain the abnormal returns (Pontiff, 1995).
Even with investors able to gain above average returns, the funds themselves usually still trade at a substantial discount. There are numerous existing explanations for the existence of closed end fund discounts/premiums outlined in the next section.

**Rational Explanations For CEF Divergence From NAV**

*Month/Seasonal*

Discounts follow season patterns with the January effect, where securities increase in value, being the most prevalent (Lee, 1990). Interestingly, the January effect is not present in the components of the mutual funds (Lee, 1990). In other words, the closed end funds are affected by seasonality while their components are not. This can be considered a rational pattern as retail investors rebalance their portfolios to reflect tax advantages.

*Agency Costs*

Funds charge an annual management fee of .5-2% of asset value (Lee, 1990). Rational investors are willing to pay less for funds with higher agency costs because the agency costs decrease investors expected returns. Open end funds have similar fees, yet they are not marked at a discount. (Lee, 1990). Additionally, research shows that the fees are non-significant in explaining the discount because funds with higher fees tended to outperform those with lower fees, and investors threaten legal action when the fees do not align with performance (Barclay, 1993).
Non Traded Securities

If a fund holds securities that are not available on the open market, NAV becomes inaccurate (Lee, 1990). This is because the stated price of an asset does not reflect the value of the asset due to thin or infrequent trading volumes. The stated price could be completely arbitrary. The arbitrary NAVs could lead to spurious recorded deviations. Most large closed end funds, however, do not hold restricted stocks, but discounts still persist (Lee, 1990).

Taxes

When a fund makes a capital gain, it has to pay capital gains taxes which can be deferred as a liability (Lee, 1990). If an investor buys shares of a closed end fund, they could be buying additional tax liabilities that are not reported in NAV (Lee, 1990). Having additional unreported liabilities is a rational explanation for the deviation from NAV.

Large Block Holders

When management owns large block of the fund, the average discount of the fund tends to be much larger than funds where management retains no ownership or very little ownership (Barclay, 1993). Research from 1979 to 1984 indicates the average discount when management owns part of the fund is 14%, while it is only 4% for funds that lack management ownership (Barclay, 1993). This observation is counter intuitive. It should be expected that when management owns part of the fund, the managers should want to liquidate the fund in order to push the value of their holdings to NAV.
Research indicates that when management holds large blocks, managers can use their voting power to secure other private benefits that are detrimental to the non-management shareholders (Barclay, 1993). In 13/19 of the management blocks analyzed, the block holders received a direct salary and in 6/19 the block holders received additional management fees (Barclay, 1993). It is worth noting that the management does not actually have to be monetary benefits, rather they only have to have the potential to use their power to receive future benefits for discounts to persist (Barclay, 1993).

Sometimes the block owners do not have to be part of management for the fund to trade at a large discount. When large nonprofits buy blocks of the fund, the chairs of the nonprofit have no material loss if the fund strays from NAV, so they have no incentive to encourage management to close the fund, therefore discounts tend to persist (Barclay, 1993).

Historical Acquisition Costs

While the bulk of the value of the CEF is explained by the fair value cost of publicly traded securities, the historical acquisition cost of the securities does play a role in the valuation of closed end funds (Carroll, 1993).

Historical cost the fund paid to acquire securities has a statistically significant effect on the fund’s equity value per share, though this effect is smaller than the current exchange value of the assets of the fund (Carroll, 1993). This is interesting because, typically, finance is forward looking, but in this case it tends to take into account backwards data. This issue could be reflective on the mispricing of the underlying assets leading to historical NAV being more
representative of the actual prices the funds could sell the assets. If the historical acquisition cost is a better reflection of the “true value” of the asset than the current market price, it is rational to consider it in the valuation of the CEF.

**Irrational Explanations For CEF Divergence From NAV**

*Days Since Inception*

When a new fund is issued at IPO, investors pay, on average, a 7% premium to acquire the shares of the fund (Lee, 1990). Research from the 1985-1987 indicates that the premium persists at a 5% rate for 20 days following the offering, then deteriorates into a an average of 10% discount by 120 days after offering (Lee, 1990).

Funds get attention during IPO, so they tend to trade at a premium, which quickly erodes. This could be due to the lack of deferred tax liability and popularization of the fund itself. Assuming that humans operate on limited rationality where information is costly to seek out, it makes sense that advertised funds should trade higher than older unpopularized counterparts.

*Noise Traders*

Closed end funds are held almost entirely by individuals rather than institutions (Lee, 1990). Individuals tend to be “Noise Traders” while institutions undergo more rigorous analysis of securities, so the discounts and premiums could represent emotions.
Using data from the 1929 period, researchers constructed the following model to use sentiment to explain the premium / discounts on closed end funds (De Long, 1990). Below is the BDL Sentiment Index and its supporting calculations:

\[
\text{BDLSentiment} = \frac{\text{S&P}}{\text{BDL}}
\]

Where

- **BDLSentiment**: An index to measure market sentiment against intrinsic values
- **S&P**: The value of the S&P
- **BDL**: The value of the BDL

\[
\text{BDL} = \frac{\text{Dt}}{\text{Rt} - \text{Gt}}
\]

Where

- **BDL**: The Barsky-De Long model of intrinsic value
- **Dt**: Last market dividend paid
- **Rt**: Last market discount rate
- **Gt**: Last dividend growth rate

Below is the scatterplot of the BDL Sentiment index vs the median premium or discount on CEFs:
Association of closed end fund discount and BDL Sentiment Index (De Long, 1990).

Figure 2 clearly shows that over the 1929-1932 period, the premiums and discounts on closed end mutual funds are positively correlated with general market sentiment.

Other research presents a similar argument that the deviations from NAV are strictly emotional. Rather than creating a new index, the other researchers simply regressed the average CEF discount against the returns of the Russell 2000 for the 1956 – 1987 period (Lee, 1991). The average discounts / premiums of all CEFs were found to be correlated with the Russel 2000 returns (Lee, 1991). The group concluded that since both the Russell and CEFs are mainly held by retail investors, retail sentiment is a primary driver in the deviation from NAV.
The fact that there are emotional investors who invest in closed end fund, could itself, explain the discounts. Noise traders increase the risk of holding the security, so rational investors demand a higher rate of compensation at all times (Lee, 1991) When funds trade at a discount they have a higher rate of return which the idea of increased compensation for noise trading risk (Thompson 1978).

Emotionally driven movements are more pronounced in closed end funds than fundamentally driven movements (Klibanoff, 1998). When a headline appeared in the New York Times regarding events that could possibly implicate the fund, the fund's movements were greater than when the fund or its components released additional financial data (Klibanoff, 1998). This again supports the view that the NAV deviation is driven by emotions. With no mechanism in the CEFs to counter the emotions of retail investors, there is no way for institutional investors to profit.
Chapter 4

Prior Research on ETF Divergence

Where the divergence of closed end mutual funds have been thoroughly researched, ETFs remain a largely unexplored area. It is important to mention that discounts and premiums do exist in ETFs, but at a much smaller magnitude. On average, funds trade at a premium of 1.1 bps (Engle, 2006). This deviation cannot be used in arbitrage situations because the bid-ask spread usually exceeds the deviation (Engle, 2006). The study provides no insight for why the premium exists, simply that, like CEFs, a deviation exists.

It is not surprising that the deviation from NAV is usually much smaller on ETFs than their CEF counterparts because ETFs can create new shares or buy back at any time. This allows the ETF to constantly try to rebalance to its NAV. Many of the factors that caused the CEF to diverge from NAV do not exist under the framework of the ETF.

In times of extreme volatility, however, ETFs can exhibit significant deviations from NAV. For example during the first weeks of August 2015, the Blackrock Dividend ETF traded at a 30% discount relative to its assets. In 2008 period, it was not uncommon for funds to trade at a 5% deviation to their net asset value, though the larger deviations tended to be centered around market opens and closes (Salsbury, 2008). Under normal circumstances, meaningful divergence is very rare, but large divergences in crisis periods are possible and persisting.
Below are the currently accepted CEF explanations for divergence when applied to the ETF.

**Rational Explanations For ETF Divergence From NAV**

*Monthly/Seasonal Changes*

Where CEFs experienced seasonality in January when discounts narrowed, there is evidence that a similar effect happens to ETFs in November (Milonas, 2008). In November, it appears that ETFs are able to track their underlining indices better than in other months. These effects are not particularly strong with the average monthly tracking error of 26bps. The reason for the tracking abilities is not explained (Milonas, 2008).

*Agency Costs*

ETFs are usually not actively managed and fees average 0.44% annually (Wall Street Journal, N.P.). While this does represent a real cost, the fees are often smaller than those of CEFs where they can range as high as 10%. Agency costs may be a factor, but they should not be as significant as CEF’s agency costs.

*Non Traded Securities*

If an ETF holds a thinly traded security, the NAV may not update often enough to be accurate. This situation represents a real problem because the NAV of the ETF may not actually represent the value of the portfolio the ETF holds. There has not been research into this area, but
from a cursory scan of ETFs many premiums / discounts can be caused by non traded securities. ETFs that track debt are especially prone to this because debt is often traded OTC, so a real market cannot be evaluated.

**Taxes**

Unlike CEFs, ETFs do not pay capital gains taxes because authorized participants officially carry all of the transactions out. This means that unpaid capital gains are not affecting the valuation of ETFs (ETF, 2016).

**Large Block Holders**

The structure of the ETF confers no benefit to large block holders, so there is no research indicating that block holders cause ETF discounts or premiums.

**Historical Acquisition Costs**

Although there is a possibility that historical acquisition costs play a role in the current valuation of the ETF this has not been studied. Logically, it should play a role in ETFs that hold non-traded security ETFs (due to the stale pricing problems), but fail to explain deviations in widely traded securities.

**Irrational Explanations For ETF Divergence From NAV**

**Days Since Inception**
ETFs do not exhibit premiums on inception like CEFs do because the creation – redemption system is immediately active. If there is a large premium or discount, the institutional investors could arbitrage it away.

*Noise Traders*

The main explanation for ETF divergence from the NAV is volatility due to noise traders. ETFs exhibit 17% more volatility than their components NAV (Cherry 2004). Excess volatility and subsequent mispricing is due to uninformed retail investors acting simultaneously (Cherry 2004). If enough “noise traders” are acting on a security at a given moment, an arbitrageur may be unable to execute at quoted prices. Additionally, the discounts from study are time varying and correlated, implying ETFs, like CEFs, move into and out of discounts / premiums in tandem (Cherry 2004).

A second study which also delves into volatility looks at the effects of ETF ownership on the securities themselves. Research indicates that as ETF ownership of an underlining security increases, the volatility of that underlining security increases (Ben-David, 2014). Specifically,

“For S&P 500 stocks, a one standard deviation change in ETF ownership is associated with a 19% standard deviation increase in intraday volatility.”

If ETFs are causing their NAV component’s volatility to increase, the quoted NAV is expected to be more inaccurate, so the possibility for a deviation is more likely.
Other Research on ETF Divergence

When a fund is diverging in value from its stated net asset value, the accuracy of the NAV needs to be scrutinized. The fund’s net asset value is calculated as

\[ \text{NAV} = \sum (\text{Portfolio Weight}) \times (\text{Last Price}) \]

This calculation assumes that the last price is reflective of the true value of the underlining securities. This is not always the case.

**Stale NAV**

Traditional open-end mutual funds have a problem where their net asset value is misstated due to stale data. (Ben-David, 2014). Other research confirms that the stale pricing problem from open ended funds carries to ETFs (Grégoire, 2011). He showed that, especially for the pre 2003 period, the discount/premium was an extraordinarily strong predictor of the movement of NAV (Grégoire 2011). That is to say if the fund was trading at a discount, the NAV would fall to eliminate the discount and the reverse would happen in the case of a premium. Post 2003 this feature was largely muted because the SEC began requiring more stringent reporting of NAV (Grégoire 2011). Nonetheless, stale pricing still represents a significant problem for the NAV.

**Big Spreads**

Like every other asset, the underlining components of an ETF have a bid-ask spread. As the bid-ask spread of those assets increases, the cost to buy or sell the entire portfolio of assets
widen. This makes it more difficult for the portfolio to be accurately represented in the stated NAV. Compounding on this issue is the fact that ETFs themselves may contribute to larger bid-ask spreads of underlining assets. When ETFs begin holding an asset, the average bid-ask spread of that asset increases (Israeli, 2015). This is because as the ETF holds an increasing percent of the float of a security, less of that security is available for market participants to trade (Israeli, 2015).

Unhinging in crisis

The net asset value for an ETF is usually strongly correlated to the exchange price of the ETF in such that they almost always move together (Milani, 2013). The correlation between price and NAV, however, loosens in the event of a crisis especially in emerging markets. This was pronounced during the Eurozone debt crisis when the correlation of Brazilian ETFs and their components fell from above .9 to below .2 (Milani, 2013). If returns in the ETF market are no longer correlated to the returns of their portfolios, then there is a real possibility for deviations to occur.

The Bank of International Settlements has taken note of the systemic risks of ETFs divergence from NAV. Specifically, they are examining the behavior of synthetic ETFs in crisis scenarios (Ramaswamy, 2011). A synthetic ETF is an ETF constructed out of derivative contracts with the intention of tracking indices. Under normal circumstances, holding a real asset is no different than holding a synthetic one, but in a crisis situation, counterparty risk becomes a concern (Ramaswamy, 2011). Under crises, a counterparty may be unable to honor their end of
the contract, so the stated NAV of the fund could be much higher than the actual value of the contracts in a synthetic ETF.
Chapter 5
Process

The purpose of this paper is to try to a better understanding as to why ETFs diverge in value from their underlining assets. This paper will attempt to explain the average premiums or discounts across a select group of ETFs rather than single ETF specific deviations. Observing averages allows for a broader market view as to when these types of anomalies occurs. Additionally, since premiums and discounts are correlated across funds, general market factors should be able to act as predictors for averages.

Based on prior research, this paper is examines when ETF deviations from intrinsic net asset value occur under the following conditions.

1. There are factors that prevent authorized participants or arbitrageurs from entering and eliminating inefficiencies
2. There is a general sense of “fear” across the markets
3. There is a significant amount of volatility or trading activity in the ETF market

Fund Selection

When selecting the funds, it is important to be mindful of the potential pitfalls of the stated net asset value. Stale pricing, international securities, and ambiguously valued ETF holdings opens up the potential for an incorrect net asset value. If the reported net asset value of
the ETF does not represent the actual value that investors can trade the underlining securities at, all reported premiums / discounts cannot be verified. All controls on the observed ETFs are intended to minimize the potential for an incorrectly state net asset value. The funds have been selected Via Bloomberg terminal’s fund screener system with controls on the following:

- **USA Domicile and USD currency**
- **Fund Asset Class Focus:** Equity
- **Fund Market Cap Focus:** Large Cap
- **Bid-Ask Spread Average over 30 Days <1%**
- **Inception Date <=1/1/2010**
- **All observed funds should have <1% spread on closing**

**USA Domicile and USD currency**

Though international funds with foreign currencies exhibit divergence from stated NAV, it may be difficult to accurately measure the value of the assets the funds hold. This is because Bloomberg records the data for end of American day rather end of international day. In essence, NAV would be inaccurate if you observe international companies which could erroneously lead to recording a divergence when the true NAV and the ETF price are actually the same. Additionally, this precaution minimizes the possibility of selecting funds with stale net asset values.

**Fund Asset Class Focus:** Equity

Equities are easier to value than debt because equities are traded on the open exchange rather than OTC in most cases. Additionally, equities ETFs do not carry the counterparty risk that synthetic ETFs have, meaning that in crisis situations equity ETFs do not have to not have uncertainty of counterparty failure which could affect the value of its holdings. Measuring only
equities ETFs ensures that the fund’s NAV is transparent, accurately stated, and carries no uncertainty of counterparty failure.

**Fund Market Cap Focus: Large Cap**

Observation has been limited to large cap tracking ETFs for a number of reasons. First, large cap stocks trade on the NYSE or the NASDAQ while small caps may trade OTC. Since NAV must be accurate, observation was limited to transparent exchanges. Second, the bid-ask spread of large-cap components tends to be small. Having smaller spreads on the underlining components is important because it removes a cost of arbitrage for the institutional investors. Finally, large cap companies have deeper market depth, meaning that large volumes of the securities can be traded at quoted prices, and institutional investors would not move markets when buying or selling the ETF components.

**Bid-Ask Spread Average over 30 Days <1%**

This is a measure of the average ask - bid divided by last trade on the ETF over the last 30 days. To be included in the sample, the value of this measure must be < 1%. The calculation accounts for all trades throughout the day, not only market opens and market closes. For examination purposes, the spread should be as low as possible so the ETFs can be accurately priced.

**Inception Date <= 1/1/2010**

Because this study is intended to observe ETFs over a number of years, it is important to have ETFs that are at least 5 years old. It may take time for a fund to mature to track its
underlining assets better as arbitrageurs become more accustomed to the price movements of the ETF relative to its underlining components.

*All observed funds should have <1% spread on closing*

Spread on closing is calculated by the average of ask–bid divided by closing price over the observation period. This additional layer of criteria is added because the spreads tend to vary over time of day. Bloomberg can only return end of day data (rather than interday), so end of day specific attributes are especially important for this examination. Since bid-ask spread end of day usually exceeds that of bid-ask spread average, all funds with large closing spreads will be removed.

**Specifications**

The time frame examined is 1/1/05-1/1/15 on a daily basis, so it can cover nearly two business cycles and have relatively up to date data. Earlier data was excluded because the SEC rule changes from 2003 that were being implemented across 2004 relating to NAV reporting (Grégoire 2011). Weekends and non trading days were not excluded to ensure absolute consistency across all funds. Excluding nontrading days causes issues when some funds use different standards to define what constitutes a nontrading day.

All data was collected through Bloomberg via functions for day end including all nontrading days and date overrides: PX_BID, PX_ASK, PX_LAST, and FUND_NET_ASSET_VAL. While it would have been ideal to pull historical minute ticks,
Bloomberg only supplies daily end of day data. It is important to note that using daily end of day data could exaggerate the general magnitude of ETF deviations from NAV because deviations most frequently occur on market opens and closes (Salsbury, 2008).

PX_BID pulls the last bid on the ETF from the date that the market is open.
PX_ASK pulls the last ask on the ETF from the date that the market is open.
PX_LAST pulls the last traded price of ETF from the date that the market is open.

FUND_NET_ASSET_VAL tracks only the stated price of a fund’s NAV based on last trade of the components that the ETF holds. This represents a stale price because it is based on last trade, rather than a real-time bid and ask price. Bloomberg does not support calculating the NAV based on real time bid-ask prices, so a proxy was used to estimate a range of what the “Real NAV” could be. The proxy is found in the Generated Fields section for a “Theoretical High NAV” and “Theoretical Low NAV”.

**Generated Fields**

Using the Bloomberg data, other fields were generated. These fields include the following:

“Bid-Ask Spread Percentage”
“Theoretical High NAV”
“Theoretical Low NAV”

“Bid-Ask Spread Percentage”
“bid-ask spread percentage” was generated to calculate how wide the bid-ask spread is for a particular ETF. This value was used to estimate how wide a spread is in terms of percent of stated value. The calculation is below.

“Bid-Ask Spread Percentage” = (PX_ASK - PX_BID)/AVERAGE(PX_ASK, PX_BID)

“Theoretical High NAV” and “Theoretical Low NAV”

While the Bloomberg generated NAV is static, a more dynamic NAV range is a better way to view the net asset value of the funds. Whenever an individual wants to trade a security there are two prices, a low bid and a high ask. NAV is only stated at closing prices without any indication of what it would actually cost to trade those underlining securities on the market at any given point. An actual market participant would only be able to replicate an ETF’s portfolio by buying each component at its ask price, and the trader would only be able to short the portfolio by selling each component at its bid price. The theoretical high/low NAV is simulating the portfolio value of the ETF as if the trader was buying it ask and selling at the bid.

Below is the method of generating a theoretical low/high NAV.

“Theoretical High NAV” = FUND_NET_ASSET_VAL + (FUND_NET_ASSET_VAL * “Bid-Ask Spread Percentage”)”

“Theoretical Low NAV” = FUND_NET_ASSET_VAL - (FUND_NET_ASSET_VAL * “Bid-Ask Spread Percentage”)
I justify using the ETF’s actual bid-ask spread percentage to create a Theoretical NAV spread because all of the assets for these funds are all liquid large cap equities. These types of equities usually have very low spreads much like the ETFs that track them. An increase in bid-asks spread of the ETF should translate to a bid-ask spread in the underlining assets.

Measuring Premiums / Discounts and Deviations

This paper analyzes and tracks six different ways to measure the funds’ deviations from NAV. The values collected represent the average premiums / discounts or deviations from NAV across the observed funds. The methods of measuring premiums and discounts are as follows:

- **Closing Premiums / Discounts**
- **Closing Deviations**
- **Bid-Ask Mid Premiums / Discounts**
- **Bid-Ask Mid Deviations**
- **Real Premiums / Discounts**
- **Real Deviations**

**Closing Premiums / Discounts**

This will take the average of each segmented fund’s (Closing Price – NAV)/NAV with extreme values of \( > \text{abs}(5\%) \) returning 0.

This measure is the traditionally stated discount or premium of the ETF relative to NAV. It however, is incomplete because it uses past prices, so the quotes could become stale. All extreme values were removed because they are often times due to unintentional misstatements in
NAV or incorrectly entered bid-ask values that do not represent and actual divergence from value.

*Closing Deviations*

This will take the average of each segmented fund’s `AbsoluteValue((Closing Price – NAV)/NAV)`. This measures the absolute value of the closing discount or premium. It can be helpful to see if the ETFs are simply deviating from their NAVs, rather than the direction of the deviation.

*Bid-Ask Mid Premiums / Discounts*

This will take the average of each segmented fund’s `(Average(Bid, Ask) – NAV)/NAV` with extreme values of > abs(5%) returning 0.

The measure will be more effective than the closing premiums / discounts to determine if the fund is trading at a discount or premium because bid-ask prices are executable, while past prices could be stale. The issue then becomes whether the bid-ask quotes are entered erroneously. All extreme values have been removed because they are often times due to unintentional misstatements in NAV or incorrectly entered bid-ask values that do not represent and actual divergence from value.

*Bid-Ask Mid Deviations*

This will take the average of each segmented fund’s `AbsoluteValue(Average(Bid, Ask) – NAV)/NAV`
The measure will show if deviations, rather than discounts or premiums happen at the same time.

"Real" Premiums/Discounts

This will take the average of each fund’s output from the following equation.

If
   BID>NAV@High, then return (BID-NAV@High)/NAV@High,
Else If
   ASK<NAV@low, then return (ASK-NAV@low)/NAV@low,
Else
   Return 0
   with extreme values of > abs(5%) returning 0.

Where

BID is the bid price of a fund at time t
ASK is the ask price of fund at time t
NAV@ High is the theoretical cost it would take to buy the underlining portfolio of securities
NAV@ Low is the theoretical cost it would take to sell the underlining portfolio of securities

Figure 3. Defining Real Premiums/Discounts
Figure 3 is a visualization of the output from the equation that determines the real premiums or discounts on an ETF. A fund can either be trading a premium (absolutely above NAV), a discount (absolutely below NAV) or at value (within the theoretical range of NAV).

Figure 4 illustrates when arbitrage is a possibility in the ETF market. The calculation for real discounts/premiums will only return a nonzero integer when arbitrage is possible. This method only marks that there is a discount or premium if the NAV falls outside of the bid ask spread. If the bid on the ETF exceeds the total value of the assets of the fund, it is classified as trading at a premium whereas, if the ask is below the value of the fund it is classified as trading at a discount. This method eliminates the real possibility of the NAV falling within the bid-ask spread and the system erroneously recording that there is an actual discount / premium. If the NAV is in between the bid-ask spread of the ETF, no arbitrage is possible, so it would not be possible to execute any transactions instantaneously to profit. Opportunities only exist when the bid-ask spread and the theoretical NAV range do not overlap because it enables investors to buy one asset and sell the other at a higher price.
"Real" Deviations

This will take the average of each funds output from the following function

AbsoluteValue of

If
  BID>NAV@High, then return (BID-NAV@High)/NAV@High,
Else If
  ASK<NAV@low, then return (ASK-NAV@low)/NAV@low,
Else
  Return 0
with extreme values of > abs(5%) returning 0.

This will measure the deviations from the NAV rather than the average of the overall premium / discount.

Developing a Regression to explain Premiums/Discounts and Deviations from NAV

After gathering the data and determining the different ways to measure discounts and premiums, I regress them against proxies for investor sentiment that could create arbitrage opportunities and factors that could inhibit sophisticated investors from executing on arbitrage opportunities. The goal is to explain what factors cause funds to trade away from their intrinsic values.

This paper’s ultimate regression will appear as follows:

“premiums/discounts and deviations from NAV”= α + βn(Explanatory variable)

Where
α is a constant

β is a coefficient on one of the explanatory variables

Defining the Explanatory Variables

There are three general themes that this paper explores.

1. There are factors that prevent authorized participants or arbitrageurs from entering and eliminating inefficiencies

2. There is a general sense of “fear” across the markets

3. There is a significant amount of volatility or trading activity in the ETF market

In the CEF market, consumer sentiment and volatility alone were shown to be factors that influenced premiums / discounts. The ETF’s creation-redemption structure allows institutional investors to counter irrational retail sentiment and volatility, thus discounts / premiums are much rarer, though they are still possible.

I predict that divergences in the ETF market only occur when there are factors limiting sophisticated investors from participating in the market, “consumer sentiment” reaches abnormal levels, and investors begin to act in large volumes at the same time. The combination of the three conditions is expected to overwhelm institutional players leading to sentiment-driven retail investors dictating the price of the ETF.

Below is the list of explanatory variables that can act as proxies for the three themes that are predicted to cause ETFs to diverge from their net asset value.

Financials Cost of Equity - Market Frictions
Dividend Yield - General Sense of Fear
ETF Average Volume - Volatility/Volume
S&P Percent Change - Volatility/Volume
Financials Cost of Equity [of the financials stocks on the S&P] - Market Frictions

With a higher cost of equity, financial firms may be unable to participate in arbitrage markets because the cost of equity increases banks’ overall financing costs. The main reason for including financials cost of equity as an explanatory variable is that it can act as a proxy for stress in the financials markets specifically. Financials are a very exaggerated cyclical industry. When the market is “down” financials decline at greater magnitude than the general market, indicating that their betas are relatively high. Research indicates that during crisis periods financials beta tends to increase (Hasnaoui, 2014) Additionally, the market risk premium of the market increases significantly because the risk free rate drops rapidly as the Federal Reserve cuts the Fed Funds target. The combination of increasing beta and MRP during a financial crisis should increase the financials cost of equity. Being the case, the cost of equity should work as a proxy for financial crisis periods.

Financial crises should be a good indicator of when firms would want to avoid participating in the ETF market’s creation-redemption mechanism. Under normal circumstances, the profit margin on the typical creation-redemption transaction is very small. During a financial crisis, firms may become unwilling to support activities that have such paper-thin margins because of general market uncertainty and possible frictions in the case of a counterparty default (resulting in failure to deliver securities agreed upon). It is predicted that as that as the probability of a financial crisis increases, less institutional investors will be willing to participate in the ETF market thus allowing deviations from net asset value to exist. Additionally, it is predicted that these deviations should be in the negative direction (discounts) because there should be more selling than buying activity on the retail side during a financial crisis.
Dividend Yield [of the stocks on the S&P] General Sense of Fear

Prior research on closed end funds used the BDLSentiment Index to estimate consumer sentiment of the marketplace. The calculation is below:

\[
\text{BDLSentiment} = \frac{\text{S&P}}{\text{BDL}}
\]

Where

BDLSentiment: An index to measure market sentiment against intrinsic values
S&P: The value of the S&P
BDL: The value of the BDL

\[
\text{BDL} = \frac{\text{Dt}}{(\text{Rt} - \text{Gt})}
\]

Where

BDL: The Barsky-De Long model of intrinsic value
Dt: Last market dividend paid
Rt: Last market discount rate
Gt: Last dividend growth rate

This effectively measured sentiment by using a combination of the S&P price with the implied S&P price using a dividend discount model. While this model could provide and effective estimate in an objective world, the Rt (Last market discount rate), is a subjective measure. To have a more objective measurement of consumer sentiment while still holding on the traditions of the model, I examined dividend yield. Dividend yield is objective and easy to collect. Additionally, is can simulate investor confidence or fear. As dividend yield rises people are more likely to be fearful of the market outlook. This is because people are discounting dividends at a higher rate implying that markets believe there is little potential for growth. The opposite is true at the lower end of dividend yield. As dividend yield falls, investors are placing more emphasis on future growing dividends rather than current payouts.

Market fear is expected to play a role in the premiums and discounts of ETFs because it leads people to make irrational decisions. When people, especially retail investors, are fearful they may move their money in or out of the ETF market without checking to see what the net
asset value of their funds are. If more investors are making decisions out of fear rather than information, it logically should lead to larger deviations from NAV. These deviations are expected to be in the negative (discount) direction because fear is associated with market downturns.

*ETF Average Volume- Volatility/Volume*

For ETF Volume, a proxy is not needed because volume can be pulled directly as it is stored historically. ETF volume is expected to cause funds to deviate from NAV because of the possibility of the overwhelming effect that it could have on institutional investors participating in the creation-redemption system. If there are more trades happening than institutional players can account for, there is bigger chance that funds will deviate from NAV. Volume should not be biased in either the positive (premium) or negative (discount) direction. If for example, there is an abnormal number of buyers on a given day, volume will be high and deviation will be positive, but if there is an abnormal number of sellers volume would also be high, but the ETFs would be trading at a discount. In both cases, a deviation from NAV occurs, but the direction is not predictable.

*S&P Percent Change- Volatility/Volume*

In “The Closed End Mutual Fund Puzzle” researchers believed that market returns are correlated with the premiums and discounts of closed end mutual funds (Lee 1991). In order to prove this, they compared the returns of the Russell 2000 to the premiums and discounts on the closed end funds (Lee 1991). They found that when the Russell 2000’s return was positive, discounts tended to be modest, but when the returns were negative discounts tended to be rather
deep (Lee 1991). They rationalized that since the closed end funds and the Russell were mostly held by retail investors, the deviations were caused by retail investor volatility (Lee 1991).

This paper is examining ETFs rather than CEFs, but the intuition of using market changes as a proxy for retail volatility remains the same. Bigger changes cause more retail investors to move money faster than institution investors can handle. The ETFs this paper examines are tracking S&P securities, so it would be inappropriate to use the Russell 2000 as the proxy. Instead, this paper uses the daily change of the S&P to track market movements. It is predicted that if the S&P’s return is positive, the ETFs should trade at a premium because investors could be rushing in and buying overvalued ETF shares, and if the S&P return is negative, retail investors could be dumping ETF shares below their NAV.
Chapter 6

Establishing that Discounts and Premiums Exist

Using the criteria laid out in the process section, I selected the following ETFs for examination. They are listed in the chart below.

Table 1. ETFs Under Examination

<table>
<thead>
<tr>
<th>QQQ</th>
<th>EEM</th>
<th>PNQI</th>
<th>IYF</th>
<th>DDM</th>
<th>DIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPY</td>
<td>QTEC</td>
<td>QLD</td>
<td>IWF</td>
<td>XLB</td>
<td>EWC</td>
</tr>
<tr>
<td>IVV</td>
<td>QQEW</td>
<td>ADRA</td>
<td>IWD</td>
<td>ERX</td>
<td>IDU</td>
</tr>
<tr>
<td>AAAXI</td>
<td>INDX</td>
<td>ADRD</td>
<td>XLK</td>
<td>RSX</td>
<td>EWG</td>
</tr>
<tr>
<td>ACWI</td>
<td>XLE</td>
<td>UPRO</td>
<td>VEA</td>
<td>XLY</td>
<td>VEU</td>
</tr>
<tr>
<td>ACWX</td>
<td>FAS</td>
<td>IWB</td>
<td>XLP</td>
<td>VGK</td>
<td>OEF</td>
</tr>
<tr>
<td>DIA</td>
<td>QQXT</td>
<td>SOXX</td>
<td>XLI</td>
<td>VIG</td>
<td></td>
</tr>
<tr>
<td>SSO</td>
<td>GDX</td>
<td>EWJ</td>
<td>XLU</td>
<td>EDC</td>
<td></td>
</tr>
<tr>
<td>XLF</td>
<td>EFA</td>
<td>ADRU</td>
<td>XLV</td>
<td>RSP</td>
<td></td>
</tr>
<tr>
<td>ADRE</td>
<td>FXI</td>
<td>EWZ</td>
<td>UYG</td>
<td>SDY</td>
<td></td>
</tr>
</tbody>
</table>

From this point forward all closing premiums / discounts, closing deviations, bid-ask mid premiums / discounts, bid-ask mid deviations, real premiums / discounts and real deviations shall refer to the arithmetic mean of the closing premiums / discounts, closing deviations, bid-ask mid premiums / discounts, bid-ask mid deviations, real premiums / discounts and real deviations of the above 56 ETFs.

Definitions and calculation process for real premiums / discounts, real deviations, closing premiums / discounts, closing deviations, bid-ask mid premiums / discounts, bid-ask mid deviations, real premiums / discounts and real deviations can be found in the process section of this paper. Below are the metrics of measuring the average premiums discounts and deviations of the ETFs under examination:
Closing Premiums / Discounts

Bid-Ask Mid Premiums / Discounts
"Real" Premiums / Discounts
Closing Deviations
Bid-Ask Mid Deviations
"Real" Deviations

Closing Premiums / Discounts

Figure 5. Average Closing Premiums / Discounts

The closing premiums / discounts are relatively small and remarkably stable over time. This implies that there may not be events where real deviations from the net asset value exist. It is important to note that the closing premiums / discounts represent trades that have already happened rather than trades that could be executed in the future. What this chart seems to prove is that the historical last trades of the ETF track the current NAV of the ETF at nearly any time \( t \).
This paper looks to examine the *current* prices an individual would be able to trade and ETF at and the *current* NAV price of the ETF, and alternative measures show a different story.

**Bid-Ask Mid Premiums / Discounts**

*Figure 6. Average Bid-Ask Mid Premiums / Discounts*

The bid-ask mid premiums / discounts measure shows a different story than its closing premiums / discounts peer. Rather than demonstrating that there are nearly no premiums or discounts overtime, the bid-ask mid data shows that during the period towards the end of 2008 to roughly 2010 ETFs could trade at an average premium or discount of up to one percent. Outside of the crisis period, however, ETFs traded very similarly to their NAV. This bid-ask mid data is significant because it shows that when you look at forward prices of ETFs (using the bid-ask prices) you find that there is a discrepancy from the current NAV.
"Real" Premiums / Discounts

Figure 7. Average Real Premiums / Discounts

Similar to the bid-ask mid premiums / discount graph, the real premiums / discounts chart shows a remarkably tight relationship between NAV and price with the exception of the 2008 to 2010 period. This data however is more valuable because it only counts real discounts / premiums that investors are able to trade off of. Since this graph will return a zero premium / discount unless the ETF bid > highest theoretical NAV price (premium) or ETF ask < lowest theoretical NAV price (discount), every time a premium / discount occurs there should be an arbitrage opportunity.

From the graph we can see that discounts and premiums persist for extended periods before reverting back to zero.

Closing Deviations
Closing deviations from NAV is measuring the average of the absolute values of the premiums and discounts found in the 56 ETFs under examination.

The closing deviations chart tells a slightly altered story from the closing discounts / premiums graph. Closing deviations, for the most part, are stable around 20 basis points except for the 2008 to 2010 period. Over that time, the closing deviations seem to be much larger. This graph, through the discrepancy with the closing premiums / discounts graph, implies that the ETFs actually do trade away from their NAV values, but without a bias in any direction (premium / discount).

*Bid-Ask Mid Deviations*
In many ways, the bid-ask mid deviations are similar to that of the closing deviations, but the magnitude is more extreme. For the vast majority of the observation period the bid-ask mid deviations from NAV was largely negligible except for the 2008 to 2010 period.

"Real" Deviations
The real deviations graph shows the same pattern as the closing deviations and the bid-ask mid deviations graph. The intensity of the pattern is greater than that of the closing deviations, but less extreme than the bid-ask mid deviations. For the observation period there does not appear to be significant deviations from the NAV aside from the late 2008 period into 2010.
Chapter 7

Single Variable Explanations

This portion of the findings section will explore the relationship between the explanatory variables outlined in the process section and the real premiums / discounts and the real deviations of the funds. This paper will not be doing the same for the closing or bid-ask mid version of the premiums / discounts or deviations measurement because this paper is examining times when real arbitrage in the ETF market may exist. Under the other measurements, a deviation could be recorded on paper, but it would be impossible to execute upon due to historical rather than forward pricing.

This chapter does not examine all the possible explanatory variables, rather it focuses on what were found to be best fits for the three themes that are predicted to cause premiums / discounts or deviations. A complete list of explanatory variables can be found in Appendix A, Appendix B, and Appendix C.

Below are the themes that are predicted to cause ETF deviations:

1. There are factors that prevent authorized participants or arbitrageurs from entering and eliminating inefficiencies

2. There is a general sense of “fear” across the markets

3. There is a significant amount of volatility or trading activity in the ETF market

The proxy for keeping authorized participants out of the market is the financials cost of equity. The financials cost of equity should do a strong job at illustrating financial stress. Since
cost of equity is calculated as Risk Free + Beta*(Market Risk Premium) and financials beta spikes during crisis, this proxy should cover stress for the sector. Financial companies are the only authorized participants in the market, so it makes sense to single them out rather than looking at a broader market.

In prior research in closed end mutual, sentiment was captured based on what was essentially the dividend yield. In the spirit of the BDL index, dividend yield was selected as a proxy for investor fear. As the yield increases, investors are more fearful and thus demand a higher discount rate for their cash flows.

The proxy for volatility is two pronged. The first part is the average volume of the selected ETFs and the second is the movement of the S&P 500. The average volume is supposed to simulate the “overwhelming effect” of ordinary investors over institutions. The higher the volume, the less able institutions will be able to correct for mispricing of the ETFs. The average movement of the S&P is added to simulate a shock to the markets. A big movement can make trading at quoted prices more difficult because they can shift before orders are executed. The larger the movement of the S&P, the more likely the ETFs are to deviate from their NAV (in the direction of the movement).

Explaining Premiums and Discounts on ETFs

This subsection uses the four explanatory variables individually to predict the premiums and discounts on the ETF market.

*Model 1: Using Financials Cost of Equity to Predict Premiums/Discounts*
### Figure 11. Average Financials Cost of Equity vs Average Premiums/Discounts Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>0.001782462</td>
<td>0.001782462</td>
<td>608.1030709</td>
<td>2.542E-124</td>
</tr>
<tr>
<td>Residual</td>
<td>3651</td>
<td>0.010701753</td>
<td>2.93118E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3652</td>
<td>0.012484215</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.004231955</td>
<td>23.96353355</td>
<td>0.003885711</td>
<td>0.000319371</td>
<td></td>
</tr>
<tr>
<td>Financials Cost of Equity</td>
<td>-0.000346956</td>
<td>-24.65974596</td>
<td>2.542E-124</td>
<td>-0.000374542</td>
<td></td>
</tr>
</tbody>
</table>

### Figure 12. Average Financials Cost of Equity vs Average Premiums/Discounts Model Output
The financials cost of equity is a moderate predictor of the actual premiums / discounts on the average ETF. The model explains 14% of the variability in premiums and discounts and appears to track the actual premiums and discounts rather well outside of the financial crisis. In the financial crisis period, however, the financials cost of equity model does not track the actual
premiums and discounts as well. This could be because the cost of equity is not updated daily thus is may run the risk of being inaccurate during that period. Nonetheless, the financials cost of equity does a good job tracking the average ETF premiums or discounts outside of the crisis period.

**Model 2: Using Dividend Yield to Predict Premiums/Discounts**

Figure 15. Average Dividend Yield vs Average Premiums/Discounts Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R                           0.287066278</td>
</tr>
<tr>
<td>R Square                             0.082407048</td>
</tr>
<tr>
<td>Adjusted R Square                    0.082155721</td>
</tr>
<tr>
<td>Standard Error                       0.001771331</td>
</tr>
<tr>
<td>Observations                         3653</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.001519898</td>
<td>9.23844E-05</td>
<td>16.45188832</td>
<td>0.001338768</td>
<td>0.001701028</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>-0.000795823</td>
<td>4.39494E-05</td>
<td>-18.10769039</td>
<td>-0.000081991</td>
<td>-0.000709655</td>
</tr>
</tbody>
</table>
Figure 16. Average Dividend Yield vs Average Premiums/Discounts Model Output

Figure 17. Average Dividend Yield vs Average Premiums/Discounts Model Output Error
The dividend yield model has a lower $r$ squared than its financials cost of equity counterpart, explaining only 8% of the variability in the premiums/discounts. The pattern of the model is rather interesting. For the financial crisis period, the model predicted that the ETFs would trade at a persisting discount. While the model undershot the magnitude of the discounts, as seen in the model output error, it still correctly predicted the direction of the discounts. This lends credence to the idea that fear can drive discounts on ETFs especially during a financial panic.

*Model 3: Using ETF Average Volume to Predict Premiums/Discounts*
Figure 19. ETF Average Volume vs Average Premiums/Discounts Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.000533356</td>
<td>5.64923E-05</td>
<td>9.441208359</td>
<td>6.37023E-21</td>
<td>0.000422596</td>
</tr>
</tbody>
</table>

Figure 20. ETF Average Volume vs Average Premiums/Discounts Model Output

Figure 21. ETF Average Volume vs Average Premiums/Discounts Model Output Error
The ETF average volume model’s output is very similar to the dividend yield output model. It appears that during the crisis period the model projects discounts, but not to the magnitude of the actual discounts. This model explains only 4% of the variability of the premiums or discounts, and it has errors clustered around the financial crisis.

**Model 4: Using S&P Percent Changes to Predict Premiums/Discounts**
Figure 23. S&P Change vs Average Premiums / Discounts Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-7.15111E-05</td>
<td>3.03142E-05</td>
<td>-2.358996059</td>
<td>-0.000130946</td>
<td>-1.20766E-05</td>
</tr>
<tr>
<td>S&amp;P Change%</td>
<td>0.023948737</td>
<td>0.002837159</td>
<td>8.44109742</td>
<td>0.018386163</td>
<td>0.029511311</td>
</tr>
</tbody>
</table>

Figure 24. S&P Change vs Average Premiums / Discounts Model Output
The S&P model to predict premiums / discounts differs significantly from its peers. Rather than showing deep discounts during the financial crisis, it shows alternating premiums and discounts. Unfortunately, this type of model does not do a good job at explaining the variability of the average premium / discounts with an r squared of a little under 2. The model maintains significance, but it is not very predictive.
Explaining Deviations on ETFs

This subsection uses the four explanatory variables individually to predict the average deviations on the ETF market.

Model 5: Using Financials Cost of Equity to Predict Deviations

Figure 27. Average Financials Cost of Equity vs Average Deviations Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.005489498</td>
<td>0.000180485</td>
<td>-30.41522745</td>
<td>-0.00584336</td>
<td>-0.005135637</td>
</tr>
<tr>
<td>Financials Cost of Equity</td>
<td>0.000564905</td>
<td>1.43793E-05</td>
<td>39.28599985</td>
<td>7.2187E-282</td>
<td>0.000536713</td>
</tr>
</tbody>
</table>
Figure 28. Average Financials Cost of Equity vs. Average Deviations Model Output

Figure 29. Average Financials Cost of Equity vs. Average Deviations Model Output Error
The financials cost of equity model does a very good job at predicting the real deviations of the ETFs from their net asset value with an r squared of almost 30%. Since this model explains so much variability of the deviations from NAV, it lends to the hypothesis that high equity costs to financial companies may stop firms from correcting deviations from NAV. It is important to note that the model is still imperfect. It is not successful at predicting the deviations during the crisis as shown by the error chart.

Model 6: Using Dividend Yield to Predict Deviations
Figure 31. Average Dividend Yield vs Average Deviations Regression Output

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.001912592</td>
<td>9.10093E-05</td>
<td>-21.01535896</td>
<td>1.18166E-92</td>
<td>-0.002091026</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>0.001716504</td>
<td>4.32952E-05</td>
<td>39.64648321</td>
<td>3.3474E-286</td>
<td>0.001631619</td>
</tr>
</tbody>
</table>

Figure 32. Average Dividend Yield vs Average Deviations Model Output
The dividend model for predicting ETF deviations from NAV is relatively strong covering 30% of all the variability. Interestingly, the dividend yield model appears to do a better job at predicting the ETF deviations from NAV in the middle of the crisis period. At the beginning of the crisis, the model overstates the deviations and towards the end, it begins to
understate it, but in the middle of the period, this model is accurate. This feature may prove valuable when this model is combined with others in a multiple regression.

**Model 7: Using ETF Average Volume to Predict Deviations**

**Figure 35. ETF Average Volume vs Average Deviations Regression Output**

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Regression 1</td>
</tr>
<tr>
<td>Residual 3651</td>
</tr>
<tr>
<td>Total 3652</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.000283712</td>
<td>5.49097E-05</td>
<td>-5.166887346</td>
<td>-0.000391369</td>
<td>-0.000176056</td>
</tr>
<tr>
<td>ETF Average Volume</td>
<td>1.34591E-10</td>
<td>3.49459E-12</td>
<td>38.51414242</td>
<td>1.2114E-272</td>
<td>1.2774E-10</td>
</tr>
</tbody>
</table>

**Figure 36. ETF Average Volume vs Average Deviations Model Output**
The ETF average volume is a relatively strong predictor of the average ETF deviation from NAV. As volume increases, the deviations also tend to increase. This is consistent with the hypothesis that additional volume could have an overwhelming effect on institutional investors that would prevent them from arbitraging all if the inefficiencies away. The model fails to predict
all periods of time with errors clustered around the 2009 crisis period. The magnitude of the errors is closer to 50bps rather than 100bps, thus making the model slightly more accurate than its peers during that time.

**Model 8: Using Absolute Value of S&P Percent Changes to Predict Deviations**

*Figure 39. Absolute Value of S&P Change vs Average Deviations Regression Output*

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.001152801</td>
<td>3.90923E-05</td>
<td>29.48916763</td>
<td>1.2548E-171</td>
<td>0.001076156</td>
</tr>
<tr>
<td>ABS: S&amp;P Change%</td>
<td>0.063200612</td>
<td>0.003659222</td>
<td>17.27159743</td>
<td>2.59346E-64</td>
<td>0.05602629</td>
</tr>
</tbody>
</table>
Figure 40. Absolute Value of S&P Change vs Average Deviations Model Output

Figure 41. Absolute Value of S&P Change vs Average Deviations Model Output Error
Figure 42. Absolute Value of S&P Change vs Average Deviations Scatterplot

The absolute value of the S&P change model to predict ETF deviations from NAV is a moderately weak model. The regression explains only around 7% of the variability in the deviations. The S&P change model is statistically significant, but not the most predictive. Like the other single variable regression models, the S&P change seems to fail to capture the magnitude of the deviations during the financial crisis.
Chapter 8
Multivariable Explanations

The last section introduced the idea that premiums, discounts and deviations could be predictable using single variable analysis. As previously hypothesized, however, it takes multiple factors to cause ETFs to deviate from the net asset value. Investors, especially retail, must have little confidence in the market, there must be violent and volatile movements, and something must inhibit institutional investors from executing arbitrage. Single variable regression would be unable to capture all the elements that make discounts and deviations possible, so the model must be expanded to account for multiple factors that influence the trading efficiency of the ETF market.

Explaining Premiums and Discounts on ETFs

Model 1: Theoretical Explanations for Premiums/Discounts plus Yesterday’s Premium/Discount

![](output)

Figure 43. Theoretical Explanations for Premiums/Discounts With Prior Day Premium/Discount Regression Output
The first multivariable model seeks to explain why ETFs on average trade at a discount or premium at a time $t$. For this model used the “financial cost of equity” as an explanation as to why
institutional investors may be unable to participate. I used dividend yield as a proxy for general market sentiment, in the sense that if yield is very high, the investors could be pessimistic about the outlook towards the future. I included the ETF average daily volume as a measure that could be used to overwhelm informed institutional investors. Finally, the daily change in the S&P was included to try to capture the direction of the average discount or premium in the market.

All of the explanatory variables are significant in the model with “yesterday's premium / discount” being the strongest. This is unsurprising as the premiums / discounts tend to be correlated day over day. All of the explanatory variables are directionally consistent with the hypothesis.

As cost of equity rises, funds tend to trade at a discount, possibly due to institutional investors additional hurdle to acquire capital. As the dividend yield rises and investors become more pessimistic of the prospect of future dividends, funds tend to trade at a deeper discount. Volume was more of a non-directional indicator. Additional volume was not included to illustrate market euphoria or pessimism, rather providing an explanation as to why the ETF market may break away from NAV. Volume is not expected to make any comment towards direction. S&P change, however, is directional and follows the theory. If the S&P is rising, ETFs are more likely to trade at a premiums, but if the S&P is falling, the funds are more likely to trade at a discount.

Overall, the model is relatively strong with a r squared of 67. The model output graph appears to mirror the actual real premiums / discounts graph. Most of the model errors tend to be centered around the 2009 period when the premiums and discounts were more extreme.

Model 2: Theoretical Explanations for Premiums / Discounts without Yesterday’s Premium / Discount regress average real discount financials cost of equity dividend yield ETF average volume S&P change
Figure 46. Theoretical Explanations for Premiums / Discounts Without Prior Day Premium / Discount

Regression Output

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs</th>
<th>F(4, 3647)</th>
<th>Prob &gt; F</th>
<th>R-squared</th>
<th>Adj R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>0.002142837</td>
<td>4</td>
<td>0.00035509</td>
<td>3,652</td>
<td>186.93</td>
<td>0.000001</td>
<td>0.1716</td>
<td>0.1707</td>
</tr>
<tr>
<td>Residual</td>
<td>0.010341236</td>
<td>3,647</td>
<td>2.9355e-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.012484072</td>
<td>3,651</td>
<td>3.4194e-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Output

| average Cardiff Discount | Coef. | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|--------------------------|-------|-----------|-------|------|---------------------|
| financials cost of equity | -0.0003214 | 0.000018 | -17.87 | 0.000 | -0.0003566 to -0.0002861 |
| dividend yield           | -0.0003637 | 0.0000525 | -5.82 | 0.000 | -0.0004081 to -0.000212 |
| etf average volume       | 2.01e-11  | 4.76e-12  | 4.22  | 0.000 | 1.08e-11 to 2.94e-11  |
| sp change                | 0.0252982  | 0.0026199 | 9.62  | 0.000 | 0.020071 to 0.0303448 |
| _const                   | 0.0043672  | 0.0001841 | 23.73 | 0.000 | 0.0040063 to 0.0047281 |

Figure 47. Theoretical Explanations for Premiums / Discounts Without Prior Day Premium / Discount Model Output

Model For Real Premium/Discount Without Yesterday
The above model seeks to predict the real premiums and discounts in the ETF market in a similar way to model 1, but without the yesterday’s premium/discount as an explanatory variable. I chose to create this model to see if all of the theoretical explanatory variables would retain their sign and if the r squared of the model would drop significantly.

The signs of all remaining explanatory variables kept the same sign as they did at the initial model. The r squared did drop off. It fell from 67 to 17. This indicates that it is very hard to predict what the premiums and discounts relative to NAV would be. Additionally, the graph output for the model seems to do a very poor job predicting the real discounts and premiums especially in the 2009 period where the model predicted only moderate discounts while the discounts were really much deeper.
Explaining Deviations on ETFs

Model 3: Theoretical Explanations for Deviations plus Yesterday’s Deviation

regress realdeviation yesterdaydeviation financialscostofequity dividendyield etfaveragevolume absspchange

Figure 49. Theoretical Explanations for Deviations With Prior Day Deviation Regression Output

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs</th>
<th>F(5, 3666)</th>
<th>Prob &gt; F</th>
<th>R-squared</th>
<th>Adj R-squared</th>
<th>Root MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>.013034294</td>
<td>5</td>
<td>.002606859</td>
<td>3,652</td>
<td>3314.59</td>
<td>0.0000</td>
<td>0.8197</td>
<td>0.8194</td>
<td>0.00089</td>
</tr>
<tr>
<td>Residual</td>
<td>.002867505</td>
<td>3,646</td>
<td>7.8648e-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.015901799</td>
<td>3,651</td>
<td>4.3555e-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| realdeviation   | Conf.     | Std. Err. | t  | P>|t| | [95% Conf. Interval] |
|-----------------|-----------|-----------|----|-----|----------------------|
| yesterdaydeviation | .8283839  | .0090875  | 91.97 | .000 | .8107237 to .8460441 |
| financialscostofequity | .0000555  | 9.91e-06  | 5.60  | .000 | .0000361 to .0000749 |
| dividendyield   | .0000928  | .0000335  | 2.77 | .006 | .0000271 to .0001594 |
| etfaveragevolume | 1.41e-12  | 2.77e-12  | 0.51  | .611 | -4.02e-12 to 6.85e-12 |
| absspchange      | .0244786  | .0018596  | 13.19 | .000 | .0208405 to .0281167 |
| _cons            | -.0007703 | .0001040  | -7.35 | .000 | -.0009759 to -.0005640 |

Figure 50. Theoretical Explanations for Deviations with Prior Day Deviation Model Output
The next model is intended to explain the average ETF real deviations from NAV. That is to say this model predicts when the average ETF is trading away from its net asset value, but makes no comment to the direction of the difference. Like the first two models in this chapter, the third model leans on a combination of factors that can be used for proxies for investor behavior as well as market data.

Yesterday’s deviation is the deviation from NAV of the prior day. It is expected that the deviations are correlated as the discounts / premiums tend to persist for a period of time. Financials cost of equity is a proxy for the general stress level of institutional investors. More stress is predicted to cause bigger deviations. Dividend yield is a proxy for investor confidence. As the yield rises, investors are more pessimistic of future earnings and are demanding immediate cash flow. Increasing pessimism should lead to larger deviations. ETF average volumes is a factor that could be used to explain the overwhelming factor of the deviations. Large volumes could flood institutional investors and cause them to be unable to correctly price the ETFs. Finally, the absolute change in the S&P is included because it could capture
violent swings. It is predicted that as the S&P moves rapidly, the ETF market will be unable to be correctly priced.

From the model output, the signs of all the variables hold. As cost of equity rises, the ETFs are unable to trade as close to their NAV. Likewise, as the dividend yield / pessimism increases, the market, again becomes less efficient. ETF average volume, though non significant (because its contribution is captured by other variables), looks like it causes larger deviations as it increases. Additionally, extreme absolute value changes of the S&P cause larger deviations.

The model output has a relatively high r squared of 81 and the graph output looks like it closely mirrors the actual deviations from NAV outlined in chapter 6. Most of the model errors are centered around the crisis period where the actual deviations were most extreme. The model does not seem to be biased during any time period with a roughly equal amount of over predictions and under predictions at any given time t.

*Model 4: Theoretical Explanations for Deviations without Yesterday’s Deviation*

\[
\text{regress realdeviation financialscostofequity dividendyield etfaveragevolume absspchange}
\]
Figure 52. Theoretical Explanations for Deviations without Prior Day Deviation Regression Output

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 3,852</th>
<th>F(4, 3647) = 611.30</th>
<th>Prob &gt; F = 0.0000</th>
<th>R-squared = 0.4014</th>
<th>Adj R-squared = 0.4007</th>
<th>Root MSE = 0.00162</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>0.00638243</td>
<td>4</td>
<td>0.001595608</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>0.003519369</td>
<td>3.647</td>
<td>2.6102e-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.015901799</td>
<td>3.651</td>
<td>4.3555e-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| realdeviation | Coef. | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|---------------|-------|-----------|-------|------|----------------------|
| financialsccostorequity | .0003071 | .0000174 | 17.69 | 0.000 | .0002731 – .0003411 |
| dividendyield | .0006614 | .0000599 | 11.04 | 0.000 | .0005439 – .0007899 |
| stfaveragevolume | 4.93e-11 | 4.96e-12 | 9.94 | 0.000 | 3.96e-11 – 5.90e-11 |
| obspricechange | .0152188 | .0033755 | 4.51 | 0.000 | .0086009 – .0218366 |
| _cons         | -.0043565 | .0001773 | -24.57 | 0.000 | -.0047041 – -.0040086 |

Figure 53. Theoretical Explanations for Deviations without Prior Day Deviation Model Output
The fourth model from this chapter seeks to explain the real deviations from NAV without using the deviation from the prior day. This model takes into account all the factors of the third model from this chapter and justifies them using the same rationale.

The signs of all the independent variables remained the same, but at the same time, the predictive power of the model declined. The model now has an $r^2$ squared of 40, significantly lower than the model which included the prior day. Additionally, the graph output shows underpredictions during the crisis period leading up to 2008, but overpredictions following the crisis resolution in 2010. Overall, this is a significantly weaker model compared with the third model in this chapter due to the inability to predict during crisis periods. The only advantage this model has is that the ETF volume is now significant because its contribution is no longer overshadowed by the discounts from the prior day.
Chapter 9

Conclusions

This paper set out to provide quantitative predictions as to when ETF premiums / discounts or deviations occur. The paper proposed that the events tended to happen under three criteria.

1. There are factors that prevent authorized participants or arbitrageurs from entering and eliminating inefficiencies
2. There is a general sense of “fear” across the markets
3. There is a significant amount of volatility or trading activity in the ETF market

Using the financials cost of equity, dividend yield, ETF volume, and S&P change as proxies for the three factors that were predicted to cause premiums / discounts and deviations, the papers constructs single and multivariable models that explain the premiums / discounts and deviations.

Table 2. Summary of Single Variable Explanations For Premiums and Discounts

<table>
<thead>
<tr>
<th>Single Variable Explanations For Premiums and Discounts</th>
<th>Adjusted R Squared</th>
<th>T Statistic on Coefficient</th>
<th>F Statistic of the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Financials Cost of Equity</td>
<td>0.14</td>
<td>-24.66</td>
<td>608.1</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>0.08</td>
<td>-18.11</td>
<td>327.89</td>
</tr>
<tr>
<td>ETF Average Volume</td>
<td>0.04</td>
<td>-12.52</td>
<td>156.9</td>
</tr>
<tr>
<td>S&amp;P Change</td>
<td>0.02</td>
<td>8.44</td>
<td>71.25</td>
</tr>
</tbody>
</table>

Table 3. Summary of Multivariable Explanations For Premiums and Discounts

<table>
<thead>
<tr>
<th>Multivariable Explanations For Premiums and Discounts</th>
<th>Adjusted R Squared</th>
<th>Weakest T Statistic on a Coefficient</th>
<th>F Statistic of the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical Explanations Plus Yesterday’s Premium/Discount</td>
<td>0.68</td>
<td>-3.31</td>
<td>1531.18</td>
</tr>
<tr>
<td>Theoretical Explanations Only</td>
<td>0.17</td>
<td>4.22</td>
<td>188.93</td>
</tr>
</tbody>
</table>
Models predicting the average premiums / discounts on the ETF market were mostly poor. The proxies for market frictions, fear, and market volatility capture only a small part of the variability in the average premiums / discounts on the ETF market. When combined in a multivariable regression, predictive ability did not improve significantly.

Table 4. Summary of Single Variable Explanations For Deviations

<table>
<thead>
<tr>
<th>Single Variable Explanations For Deviations</th>
<th>Adjusted R Squared</th>
<th>T Statistic on Coefficient</th>
<th>F Statistic of the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Financials Cost of Equity</td>
<td>0.30</td>
<td>39.29</td>
<td>1543.39</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>0.30</td>
<td>39.64</td>
<td>1571.84</td>
</tr>
<tr>
<td>ETF Average Volume</td>
<td>0.29</td>
<td>38.5</td>
<td>1483.34</td>
</tr>
<tr>
<td>S&amp;P Change</td>
<td>0.08</td>
<td>17.27</td>
<td>298.308</td>
</tr>
</tbody>
</table>

Table 5. Summary of Multivariable Explanations For Deviations

<table>
<thead>
<tr>
<th>Multivariable Explanations For Deviations</th>
<th>Adjusted R Squared</th>
<th>Weakest T Statistic on a Coefficient</th>
<th>F Statistic of the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical Explanations Plus Yesterday's Deviation</td>
<td>0.82</td>
<td>0.51</td>
<td>3314.59</td>
</tr>
<tr>
<td>Theoretical Explanations Only</td>
<td>0.40</td>
<td>4.51</td>
<td>611.30</td>
</tr>
</tbody>
</table>

Models predicting average deviations on the ETF market seemed to fare significantly better. Both the single and multivariable models were able to capture over 30% of the variability on deviations with notable shortcomings during crisis periods. The discrepancy between the predictive power of the average premiums / discounts vs the average deviations models implies that it is possible to estimate when funds will break away from its NAV, but not the direction. That is to say, it may not be possible to estimate when the ETFs will be trading at a premium or a discount, but it is possible to model when the ETF’s exchange value differs from its net assets.

When predicting for deviations, rather than premiums / discounts, all three criteria (through their proxies) are statistically significant. In other words, when institutional players face friction, there is general market fear, and volume/volatility in the market is high, ETFs are more likely to trade away from their stated net asset value.
Appendix A

Other Explanatory Variables

This section explored the explanatory variables that were not used in the process section of this paper. Each explanatory variable represents a potential proxy for market frictions, investor fear, or volume/volatility. Below are the variables that were not used in the process and single variable regression section:

- Average Bid/Ask Spread % [of the Stocks listed on the S and P]
- Average WACC
- Average D/E
- Average Short Interest
- Median P/E
- Average P/E
- Average P/B
- Average Daily Volume
- Financials WACC
- Financial Cost of Debt
- Financials Cost of Equity
- Financials Beta
- S&P Vol
- Dow Vol
- Russell Vol
- Institutional Ownership
- Insider Ownership
- Dividend Yield
- Average EPS
- S&P, Dow, and Russell Change%
- VIX Index
- Credit Suisse Fear Barometer
- Fed Funds
- Dollar Strength
- Barclays Junk Index
- Ibox Investment Grade Index
- Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer Confidence Index
- Gold, Silver Copper, and Oil
- ETF Average Volume
- Liquidity Component to Cleveland Fed Financial Stress Index
Cleveland Fed Financial Stress
Market Crash Portion of Cleveland Fed Financial Stress Index
Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index
Interbank or Funding Markets Portion of Cleveland Financial Stress index
St Louis Fed Stress Index

Average Bid/Ask spread % [of the stocks listed on the S&P]
As the bid-ask spread widens less market makers are participating. As less investors are willing to participate in a market, sentiment is assumed to become worse. Wider spreads represent a cost to arbitrageurs as they must cross the spread to compete their transaction. It is predicted that as the spread widens, the deviations from NAV become more extreme.

Average WACC [of the stocks listed on the S&P]
A higher WACC could prevent institutional investors from taking advantages of arbitrage because they carry a higher cost of capital. Institutions could chose to allocate funds into higher returning segments compared to relatively small opportunities in the ETF market.

Average D/E [of the stocks listed on the S&P]
This is a measure of leverage across all of the firms in the S&P 500. The idea is that highly levered firms may be unable to finance functions related to capturing arbitrage in the ETF market. Unnecessary leverage could represent a market friction.

Average Short Interest [of the stocks listed on the S&P]
This measures the average percent of float in an S&P 500 company that originates from investors shorting. An elevated level of short interest indicates that markets may not confident on
evaluations. Less confidence could lead retail investors to begin pulling out of ETFs thus causing them to trade at a discount.

*Median P/E and Average P/E [of the stocks listed on the S&P]*

With a higher P/E it is assumed that investors are more confident of the growth of institutions. This opens the opportunity for manias to be more likely, but at the same time could mean that enough investors are willing to participate in the markets to eliminate arbitrage opportunities.

*Average P/B [of the stocks listed on the S&P]*

As P/B rises, investors are typically placing more confidence on the earnings potential of a company rather than its physical assets minus liabilities. This implies investor confidence and smaller deviations from net asset value.

*Average Daily Volume [of the stocks listed on the S&P]*

This explanatory variable will measure the trading volume of the average S&P 500 company on a given day. This could be a powerful explanatory variable for deviations from intrinsic value because if a large group of investors shifts their money rapidly, institutional players may be unable to correct for market inefficiencies due to limited budgets.

*Financials WACC [of the financials stocks listed on the S&P]*

All of the preferred dealers for ETFs are financials companies, so this variable can explain when their cost of capital is increased. In the same way that WACC could be used to
explain a cost of arbitrage, Financials WACC focuses specifically on the cost of financials ability to arbitrage.

*Financial Cost of Debt [of the financials stocks listed on the S&P]*

Financial institutions are heavily funded though debt. When the cost of debt is more expensive, they may not be able to allocate as many resources to capturing arbitrage opportunities on the ETF market.

*Financials Cost of Equity [of the financials stocks listed on the S&P]*

With a higher cost of equity, financial firms may be unable to participate in arbitrage markets because the cost of equity increases banks’ overall financing costs. A higher cost of equity is expected to produce larger deviations on the ETF market.

*Financials Beta [of the financials stocks listed on the S&P]*

Beta is a measure of volatility against the S&P 500. If the financials volatility is increasing relative to the rest of the market, that could signify a problem with the companies. Any possibility of a financial freeze could halt the ability of banks to participate in the ETF market.

*S&P Volatility*

If the markets are generally volatile, there is the money may be moving in and out of the market rather quickly. If the money is moving faster than the institutional investors can account for, there may be deviations from NAV.
**Dow Volatility**

The same justification for examining the volatility of the S&P applies to examining the volatility of the Dow Jones Industrial Index. More volatility is expected to induce more deviations from NAV.

**Russell Vol**

Again, measuring the volatility of the Russell 2000 could offer insights as to when money is moving in and out of the financial markets. The more movement, the more likely that ETF will trade away from their stated net asset value.

**Institutional Ownership [of the stocks listed on the S&P]**

When institutional ownership of S&P stocks is higher, that indicates that money may not move as quickly as if it was owned by retail investors. Institutional owners are more likely to have the tools to correct for and profit off of market inefficiencies. With less institutional owners, there is an increased possibility in fund divergence.

**Insider Ownership [of the stocks listed on the S&P]**

Like institutional owners, insiders are predicted to be more knowledgeable than the typical retail investor. The absence of insider owners in the overall S&P 500 could indicate that less informed money is deciding the ETF market’s price.

**Dividend Yield [of the stocks listed on the S&P]**
When dividend yield falls that could indicate a number of things, but for the purposes of market sentiment, a low yield could mean that investors are more confident about future earnings than present ones. As confidence rises, the deviations from navel ought to be smaller or the funds should trade at a premium as retail investors are more willing to buy into the ETF market.

*Average EPS [of the stocks listed on the S&P]*

A low average EPS shows that the average earnings of an S&P 500 company are either falling or being diluted by new shares. In either case, this could indicate troubled markets that would inhibit institutional participation in the ETF market.

*S&P, Dow, and Russell Change%*

When the major indexes move significantly actors in the market are adding or withdrawing a significant amount of money. When the move is large enough, institutional investors could be taken by surprise and inefficiencies could appear. Thus, larger changes are expected to lead to bigger deviations from net asset value.

*VIX Index*

The VIX index is often coined as the fear index. It is the expected market volatility based on the options pricing. More fear could lead to ETFs diverging from their net asset value.

*Credit Suisse Fear Barometer*

The Fear Barometer is similar to the VIX in that is generated by the implied volatility using options, but it is different in that it only accounts for implied volatility from short
positioned products. This is to say that it is only accounting for downside volatility of the financial markets. This could be a useful tool to estimate consumer sentiment.

*Fed Funds*

The Fed Funds rate is the benchmark interest rate for most banking institutions. If the Fed Funds rate is “too high” or moves unexpectedly, institutional investors may not be able to participate in arbitrage activities.

*Dollar Strength*

A very strong dollar could indicate a shortage of dollars in the financial markets. Increasing strength may lead to institutional investors participating less in the arbitrage of ETF markets.

*Barclays Junk Index*

The Junk Index tracks the performance of “junk” rated bonds. If junk bonds are performing worse, it may be a sign that there is trouble in the financial markets. More trouble is expected to lead ETFs to wider deviations from their NAV.

*Ibox Investment Grade Index*

The investment grade index tracks the performance of investment grade bonds. If investment grade bonds are doing poorly, it could be an indication of financial crisis that would stop institutional participation in the ETF markets.
Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer confidence index

All three indexes are generated from polls for consumer confidence. Ideally, these can be considered proxies to measure the general confidence level of the consumers. More confidence is expected to produce lower deviations from NAV.

Gold, Silver Copper, and Oil

Commodities are included into the examination because large movements could indicate real financial panics.

ETF Average Volume

If ETF volume is high, many people are trading ETFs. If too many individual investors are trading ETFs, they could force the price out of the efficient range, thus causing a divergence.

% Retail Owned [of the stocks listed on the S&P]

Retail investors tend to be less informed than their institutional peers. As retail ownership rises, I expect there to be an elevated chance of fund divergence, because there will be less players attempting to execute on arbitrage opportunities.

% Retail owned [of the observed ETFs]

Retail investors tend to be less informed than their institutional peers. As retail ownership rises, I expect there to be an elevated chance of fund divergence, because there will be less players attempting to execute on arbitrage opportunities.
**Liquidity Component to Cleveland Fed Financial Stress Index**

The Liquidity Component to Cleveland Fed Financial Stress Index calculates the spread between bank debt and US government benchmark rates. A wider spread indicates there could be liquidity problems in the market place. Liquidity problems may inhibit market participants from allocating funds to execute on ETF arbitrage.

**Cleveland Fed Financial Stress**

This index is generated to track the overall stress of the financial sector. It accounts for liquidity, market crashes, volatility and other factors. The idea that this could explain ETF divergence from NAV is rooted in the ideal that under financial stress, firms may be unable to participate in the ETF market to correct for mispricing.

**Market Crash Portion of Cleveland Fed Financial Stress Index**

This is a component of financial stress as defined by the Cleveland Fed. If the market is “crashing”, banks may be less willing to participate in the ETF marketplace to drive out inefficiencies.

**Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index**

This measures the spread between the Libor and the Fed Funds rate. When the spread widens, it is implied financial stress because the actual bank borrow rate is diverging from what the Fed is targeting. More spread indicates market participants may not be able to fund participation in the ETF market.
Interbank or Funding Markets Portion of Cleveland Financial Stress Index

This measures the interbank liquidity spread and spreads of various levels of corporate bonds. Ideally, it should be a proxy of the premiums banks pay for funds. When this premium is higher, banks may have trouble executing arbitrage opportunities in the ETF markets.

St Louis Fed Stress Index

Like the Cleveland Fed Stress index, the St. Louis Fed Stress index tracks the general financial stress of the economy. It is updated less frequently and uses slightly altered criteria.
Appendix B

Other Single Variable Explanations for Real Premiums / Discounts

Appendix B includes all explanatory variables that attempt to predict the real premiums / discounts in the ETF market. This section also includes commentary on the relationship of each of the variables with real premiums / discounts.

Single Variable Regression Table Explaining “Real Premiums / Discounts” Ordered By R Squared

<table>
<thead>
<tr>
<th>Explanatory Variables For &quot;Real Premium/Discount&quot;</th>
<th>R Squared</th>
<th>Coefficient</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yesterday Discount</td>
<td>0.63859335</td>
<td>0.79916</td>
<td>-1.4E-05</td>
</tr>
<tr>
<td>Financials Cost of Equity</td>
<td>0.14277726</td>
<td>-0.00035</td>
<td>0.004232</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>0.08240705</td>
<td>-0.0008</td>
<td>0.00152</td>
</tr>
<tr>
<td>Conference Board Consumer Confidence Index</td>
<td>0.08117073</td>
<td>2.33E-05</td>
<td>-0.00182</td>
</tr>
<tr>
<td>Bloomberg Confidence</td>
<td>0.06427141</td>
<td>6.06E-05</td>
<td>-0.00215</td>
</tr>
<tr>
<td>Fed Funds</td>
<td>0.06327556</td>
<td>0.000231</td>
<td>-0.00043</td>
</tr>
<tr>
<td>Barclays Junk Index</td>
<td>0.05653767</td>
<td>-0.00014</td>
<td>0.000748</td>
</tr>
<tr>
<td>Median P/E</td>
<td>0.05343734</td>
<td>0.00189</td>
<td>-0.0033</td>
</tr>
<tr>
<td>Liquidity Component to Cleveland Fed Financial Stress Index</td>
<td>0.05310011</td>
<td>-0.00058</td>
<td>0.001517</td>
</tr>
<tr>
<td>Financials Beta</td>
<td>0.04951644</td>
<td>-0.0023</td>
<td>0.00277</td>
</tr>
<tr>
<td>ETF Average Volume</td>
<td>0.04120274</td>
<td>-4.5E-11</td>
<td>0.000533</td>
</tr>
<tr>
<td>Average Daily Volume</td>
<td>0.03648193</td>
<td>-1.7E-10</td>
<td>0.000929</td>
</tr>
<tr>
<td>Michigan Confidence</td>
<td>0.03641041</td>
<td>3.43E-05</td>
<td>-0.00271</td>
</tr>
<tr>
<td>Change in VIX Index %</td>
<td>0.02630784</td>
<td>-0.00501</td>
<td>-5.8E-05</td>
</tr>
<tr>
<td>VIX Index</td>
<td>0.0256328</td>
<td>-3E-05</td>
<td>0.000525</td>
</tr>
<tr>
<td>Copper</td>
<td>0.02379108</td>
<td>1.7E-07</td>
<td>-0.00123</td>
</tr>
<tr>
<td>Financial Cost of Debt</td>
<td>0.02216627</td>
<td>0.000312</td>
<td>-0.00084</td>
</tr>
<tr>
<td>Financial Cost of Debt Change%</td>
<td>0.01914735</td>
<td>0.023949</td>
<td>-7.2E-05</td>
</tr>
<tr>
<td>S&amp;P Change%</td>
<td>0.01914735</td>
<td>0.023949</td>
<td>-7.2E-05</td>
</tr>
<tr>
<td>Average Debt/Equity</td>
<td>0.01859109</td>
<td>-5.4E-06</td>
<td>0.000806</td>
</tr>
<tr>
<td>Financials Cost of Equity Change%</td>
<td>0.01806708</td>
<td>0.025387</td>
<td>-7.1E-05</td>
</tr>
<tr>
<td>Dow Change%</td>
<td>0.01806708</td>
<td>0.025387</td>
<td>-7.1E-05</td>
</tr>
<tr>
<td>ABS: ETF Volume Change</td>
<td>0.01788538</td>
<td>-2.7E-11</td>
<td>0.000179</td>
</tr>
<tr>
<td>St Louis Fed Stress Index</td>
<td>0.01768502</td>
<td>-0.00019</td>
<td>-0.00012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Russell Vol</td>
<td>0.01742461</td>
<td>-1.8E-05</td>
<td>0.000351</td>
</tr>
<tr>
<td>Financials Beta Change%</td>
<td>0.0168061</td>
<td>0.017396</td>
<td>-7.1E-05</td>
</tr>
<tr>
<td>Russell Change%</td>
<td>0.0168061</td>
<td>0.017396</td>
<td>-7.1E-05</td>
</tr>
<tr>
<td>Median P/E Change in %</td>
<td>0.01488439</td>
<td>0.020142</td>
<td>-6.8E-05</td>
</tr>
<tr>
<td>Average Bid - Ask Spread %</td>
<td>0.01300055</td>
<td>0.008749</td>
<td>-0.00017</td>
</tr>
<tr>
<td>Credit Suisse Fear Barometer</td>
<td>0.01265164</td>
<td>3.75E-05</td>
<td>-0.00092</td>
</tr>
<tr>
<td>S&amp;P Vol</td>
<td>0.01114771</td>
<td>-1.6E-05</td>
<td>0.000211</td>
</tr>
<tr>
<td>Oil</td>
<td>0.01042055</td>
<td>9.41E-06</td>
<td>-0.00083</td>
</tr>
<tr>
<td>Dow Vol</td>
<td>0.00993701</td>
<td>-1.7E-05</td>
<td>0.000202</td>
</tr>
<tr>
<td>Average Short Interest</td>
<td>0.00992037</td>
<td>0.000286</td>
<td>-0.00116</td>
</tr>
<tr>
<td>Market Crash Portion of Cleveland Fed Financial Stress Index</td>
<td>0.0098092</td>
<td>-2.7E-05</td>
<td>0.000278</td>
</tr>
<tr>
<td>Cleveland Fed Financial Stress</td>
<td>0.00965742</td>
<td>-0.00016</td>
<td>-1.8E-05</td>
</tr>
<tr>
<td>ETF Volume Change</td>
<td>0.0093633</td>
<td>1.36E-11</td>
<td>-7.1E-05</td>
</tr>
<tr>
<td>Change in Credit Suisse Fear Barometer %</td>
<td>0.00885951</td>
<td>0.005794</td>
<td>-7E-05</td>
</tr>
<tr>
<td>Yesterday Deviation</td>
<td>0.00883564</td>
<td>0.122609</td>
<td>-0.00037</td>
</tr>
<tr>
<td>Change in Ibox Investment Grade Index %</td>
<td>0.00686683</td>
<td>-0.0546</td>
<td>-5.8E-05</td>
</tr>
<tr>
<td>Average WACC</td>
<td>0.00525963</td>
<td>-0.00025</td>
<td>0.002222</td>
</tr>
<tr>
<td>Change in Conference Board Consumer confidence index %</td>
<td>0.00469534</td>
<td>-0.0063</td>
<td>-6.6E-05</td>
</tr>
<tr>
<td>Interbank or Funding Markets Portion of Cleveland Financial Stress index</td>
<td>0.00380465</td>
<td>-4.1E-05</td>
<td>0.000171</td>
</tr>
<tr>
<td>Dividend Yield Change in %</td>
<td>0.00330767</td>
<td>-0.00626</td>
<td>-6.5E-05</td>
</tr>
<tr>
<td>Gold</td>
<td>0.00233206</td>
<td>-2.2E-07</td>
<td>0.000169</td>
</tr>
<tr>
<td>Change in Silver %</td>
<td>0.00221083</td>
<td>0.004711</td>
<td>-6.9E-05</td>
</tr>
<tr>
<td>Change in Michigan Confidence %</td>
<td>0.00215549</td>
<td>-0.00791</td>
<td>-6.6E-05</td>
</tr>
<tr>
<td>Institutional Ownership</td>
<td>0.00203402</td>
<td>-2E-05</td>
<td>0.001825</td>
</tr>
<tr>
<td>% Retail owned (S&amp;P)</td>
<td>0.00197283</td>
<td>2.01E-05</td>
<td>-0.00012</td>
</tr>
<tr>
<td>Change in Gold %</td>
<td>0.00186965</td>
<td>0.007602</td>
<td>-6.9E-05</td>
</tr>
<tr>
<td>Change in Interbank or Funding Markets Portion of Cleveland Financial Stress index</td>
<td>0.00181948</td>
<td>0.007575</td>
<td>-6.6E-05</td>
</tr>
<tr>
<td>Dollar Strength</td>
<td>0.00180439</td>
<td>1.77E-05</td>
<td>-0.00151</td>
</tr>
<tr>
<td>Retail ETF Ownership</td>
<td>0.00150489</td>
<td>5.36E-06</td>
<td>-0.00023</td>
</tr>
<tr>
<td>Institutional ETF Ownership</td>
<td>0.00150489</td>
<td>-5.4E-06</td>
<td>0.000303</td>
</tr>
<tr>
<td>Change In Market Crash Portion of Cleveland Fed Financial Stress Index</td>
<td>0.00136941</td>
<td>-0.0008</td>
<td>-6.4E-05</td>
</tr>
<tr>
<td>Average P/B</td>
<td>0.00133139</td>
<td>7.81E-06</td>
<td>-0.00013</td>
</tr>
<tr>
<td>Average EPS</td>
<td>0.00128077</td>
<td>3.18E-05</td>
<td>-0.00011</td>
</tr>
<tr>
<td>Insider Ownership</td>
<td>0.00122709</td>
<td>0.000925</td>
<td>-0.0006</td>
</tr>
<tr>
<td>Average P/E Change in %</td>
<td>0.0011475</td>
<td>0.001531</td>
<td>-6.8E-05</td>
</tr>
<tr>
<td>Change in Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index</td>
<td>0.00114524</td>
<td>0.0006</td>
<td>-6.8E-05</td>
</tr>
<tr>
<td>Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index</td>
<td>0.00113149</td>
<td>-6.2E-05</td>
<td>1.67E-05</td>
</tr>
<tr>
<td>ABS: Dow Change%</td>
<td>0.00105907</td>
<td>-0.00726</td>
<td>-2.9E-05</td>
</tr>
</tbody>
</table>
Explanatory Variables for “Real Premiums / Discounts” and Real Deviations ordered as observed:

Average Bid/Ask Spread % [of the Stocks listed on the S and P]
Average WACC
Average D/E
Average Short Interest
Median P/E
Average P/E
Average P/B
Average Daily Volume
Financials WACC
Financial Cost of Debt
Financials Cost of Equity
Financials Beta
S&P Vol
Dow Vol
Russell Vol
Institutional Ownership
Insider Ownership
Dividend Yield
Average EPS
S&P, Dow, and Russell Change%
VIX Index
Credit Suisse Fear Barometer
Fed Funds
Dollar Strength
Barclays Junk Index
Ibox Investment Grade Index
Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer Confidence Index
Gold, Silver Copper, and Oil
ETF Average Volume
Liquidity Component to Cleveland Fed Financial Stress Index
Cleveland Fed Financial Stress
Market Crash Portion of Cleveland Fed Financial Stress Index
Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index
Interbank or Funding Markets Portion of Cleveland Financial Stress index
St Louis Fed Stress Index

Yesterday Premium/Discount
Yesterday’s discount / premium does a good job at explaining the current discount or premium as shown by the chart. This is expected because from the graph of the real premium or discount overtime, it appears that the average remains similar to the prior days’ before reverting to zero.

*Average Bid/Ask Spread % [of the Stocks listed on the S and P] v Real Premium / Discount*

![Graph of Yesterday Premium/Discount](image)

![Graph of Average Bid - Ask Spread %](image)
Both the average bid-ask spread and the daily change in the average bid/ask spread of the S&P 500 companies does a poor job predicting what the real premium and discount would be. Although the relationship is slightly positive, it is not enough to justify further examination.

Average WACC v Real Premium / Discount
Both the average WACC and the average graph in WACC do a poor job at predicting the average real premium or discount of the ETFs. It was expected that as the cost of capital increases, or changes rapidly, there would be bigger premiums or discounts, but this is not the case.

\[ \text{Average D/E v Real Premium / Discount} \]
Though a slightly better explanation than average WACC at explaining discounts and premiums, this measure fails to capture much meaningful variability in average real discounts and premiums of the ETF market. This relationship will not be examined further.

*Average Short Interest v Real Premium / Discount*
I had a strong feeling that short interest would be a good proxy for sentiment of investors in the market. Unfortunately, short interest does an extremely poor job explaining the average real premiums or discounts in the ETF market. Short interest warrants no further examination.

*Median P/E v Real Premium / Discount*
Median P/E seems to have a positive relationship with the average premium or discount of the ETF market. Previously, this paper predicted that P/E could be used for a proxy for investor sentiment. As it rises the sentiment is expected to increase and as sentiment increases, it could be predicted that ETF trade a premium. This is consistent with the median P/E chart. It is

\[ R^2 = 0.0534 \]

\[ R^2 = 0.0149 \]
important to note that the relationship is not particularly strong, accounting for only around 5% of the variability.

For the median P/E change, the relationship seems to be more of a random scatter, so it will not be examined further.

**Average P/E v Real Premium / Discount**

*Average P/E*

![Graph](image)

*Average P/E Change in %*

![Graph](image)
Unlike median P/E average P/E is heavily skewed by individual companies, so the data really is not representative of an “average firm”. This data is not going to be examined further.

Average P/B v Real Premium / Discount

Average P/B

Average P/B Change in %
Like average price to earnings, price to book is skewed heavily by individual companies, so this data is not useful as an explanatory variable.

**Average Daily Volume v Real Premium / Discount**

*Graph showing the relationship between average daily volume and real premium/discount.*

*R² = 0.0365*

*Graph showing the relationship between average daily volume change in % and real premium/discount.*

*R² = 0.0005*
The average daily volume chart provides an interesting pattern. It appears that when volume in the S&P is low there does not appear to be large discounts or premiums in the ETF markets. Larger volumes appear to lend themselves to larger premiums or discounts. Unfortunately, the chart does not do a good at predicting the direction of the premium or discount. This could offer some support to the idea that large volumes of trades could overwhelm institutional investors thus leading to premiums or discounts.

The change in volume does not do a good job at explaining the premiums or discounts on ETFs.

*Financials WACC v Real Premium / Discount*
Both the financials WACC and the financials change in WACC fail to explain a significant portion in the variability in real premiums or discount so ETFs. These relationships will not be examined further.

Financial Cost of Debt v Real Premium / Discount
Like financials WACC, financials cost of debt fails to explain premiums and discounts.

Financials Cost of Equity v Real Premium / Discount
Financials cost of equity does a good job at explaining real premiums or discounts on ETFs. It appears that as the financials cost of equity increases, the funds tend to trade at a discount. This could be because as banks equity costs more, they become unwilling to participate in the ETF market, thus opening the door for deep discounts in ETFs.

The financials cost of equity change does not appear to have any meaningful effect on the premiums or the discounts so it will not be examined further.

Financials Beta v Real Premium / Discount
Financials beta appears to have a weak negative relationship with premiums and discounts. As financials beta increases, the ETF tend to trade at a discount. The problem, however, appears to be that the relationship fans out when the betas get extreme. The relationship may be explained by the idea that as financials are becoming risker investments, the companies may not be able to participate in the ETF market.
Financials beta change % produces random scatter and will not be examined further.

S&P Vol v Real Premium / Discount

S&P volatility produces an interesting image. It does not appear scattered, rather when volatility is low, there is not a premium or discount. When volatility is above 20 the data fans
out. This is not a good explanatory variable because of the fanning, but the fanning may provide a good explanation for the divergence from the ETF NAV.

*Dow Vol v Real Premium / Discount*

![Dow Vol v Real Premium / Discount](image)

![Dow Vol Change in %](image)
Like the S&P volatility, the Dow volatility is not necessarily a good explanation variable, but it provides a pattern that may provide insights to the deviations from NAV, but not necessarily the direction of deviation.
Like the S&P volatility, the Russell volatility is not necessarily a good explanation variable, but it provides a pattern that may be examined later.

**Institutional Ownership v Real Premium / Discount**

![Institutional Ownership plot]

**Insider Ownership v Real Premium / Discount**

![Insider Ownership plot]
Prior research indicated heavily that retail ownership (and conversely low institutional ownership) is one of the causes of fund deviations from NAV. This graph, however, does not support that assumption at all. Instead, the data appears as random scatter.

**Dividend Yield v Real Premium / Discount**
When first included as a variable to explore, I did not think that average dividend yield of the S&P would be a good explanatory variable for ETF premiums and discounts. The idea was that it could be used for a proxy of investor sentiment. As yield falls, investors are looking less for safety and more for risk indicating they are confident in the market. As it rises, investors are being more pessimistic and are demanding higher yields to invest. The data appears consistent with the idea that sentiment drives discounts and premiums. Therefore, when the yield increases it appears that ETFs are more likely to trade at a discount.

The yield change however, appears as a random scatter.

*Average EPS v Real Premium / Discount*

![Graph showing average EPS vs. discount/premium change]  

Average EPS is too skewed to be considered a good explanatory variable.

*S&P, Dow, and Russell Change% v Real Premium / Discount*
S&P Change%  
\[ R^2 = 0.0191 \]

Dow Change%  
\[ R^2 = 0.0181 \]
All S&P, Dow, and Russel changes do not explain the real premiums and discounts of the ETFs. They all appear to be more random scatter. This is disappointing because one hypothesis was that big changes could be one of the drivers of ETF premiums or discounts in the direction of the change.
The VIX does not provide the best explanation for the premiums or discounts on ETFs, but it does appear to be more than just random scatter. As the VIX is low, the premiums and discounts are not extreme, but as VIX increases the deviations seem to happen.
The Fear Barometer seems to show random scatter.

Fed Funds v Real Premium / Discount
Both the Fed Funds rate and the change in Fed Funds rate show random scatter. They provide little insight into whether the ETFs are trading at a discount or premium.

*Dollar Strength v Real Premium / Discount*
Dollar Strength also does not explain the premiums or discounts in the ETF market, but instead shows random scatter.

*Barclays Junk Index v Real Premium / Discount*
Barclays Junk index appears to do an average job at predicting the premiums or discounts in the ETF market. Unfortunately, the relationship is opposite of what was expected. It was anticipated that as junk does poorly, there would be discounts on the ETF market (because market sentiment should be low). Instead, as junk does poorly there appears to be premiums.

The change in the Junk index provides random scatter.
The Investment Grade index provides random scatter and will not be examined further.
Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer Confidence Index v Real Premium / Discount

**Michigan Confidence**

- **R² = 0.0358**
- Range: -2.00% to 1.00%
- Years: 45 to 95

**Bloomberg Confidence**

- **R² = 0.0643**
- Range: -2.00% to 1.00%
- Years: 25 to 50
All of the consumer sentiment measures appear to produce the same almost random scatter pattern. The only interesting parts of the graph appear at the lower ends. As consumer confidence deteriorates, it appears that discounts and premiums become extreme. Additionally, it can be inferred that as consumer sentiment is low, deviations tend to be larger and negative more often.

*Gold, Silver Copper, and Oil v Real Premium / Discount*
Silver

$R^2 = 3 \times 10^{-7}$

Discount/Premium

Change in Silver %

$R^2 = 0.0022$
Copper

\[ R^2 = 0.0238 \]

Change in Copper %

\[ R^2 = 7 \times 10^{-6} \]
All of the commodities under examination fail to provide meaningful explanation for the variability in the ETF discounts or premiums.

*ETF Average Volume v Real Premium / Discount*
ETF volume seems to be an average explanatory variable for premiums or discounts on the ETF market. As ETF volume increases, the funds are more likely to trade at discount. This could be because retail investors are driving money out of the ETF market faster than institutions can correct for mispricing. The correlation is tighter at less extreme volumes.

ETF volume change is not a good explanatory variable because of a low r squared, but it shows a clear pattern. It looks like when the volume is less extreme, the discounts and premiums are less extreme, but when the volume change is more extreme, the discounts and premiums become bigger. This could be because this variable is may be better at explaining deviations rather than premiums and discounts.

*Liquidity Component to Cleveland Fed Financial Stress Index v Real Premium / Discount*
As the Liquidity Component to the Cleveland Fed Stress index rises, it appears that ETFs become increasingly discounted. It is not a purely linear relationship. The discounts start to appear after the index passes 3. After that point, it does not seem to matter what the index is at; all the discounts are similar.

_Cleveland Fed Financial Stress v Real Premium / Discount_
While the Liquidity Portion of Financial Stress was an ok explanation, the general financial stress is a less strong explanatory variable. Since it appears to be close to random scatter, I will not be examining it further.
Discount

Market Crash Portion of Cleveland Fed Financial Stress Index

\[ R^2 = 0.0098 \]

Change In Market Crash Portion of Cleveland Fed Financial Stress Index

\[ R^2 = 0.0014 \]
The Market Crash Portion of the Cleveland Fed index also does not do a good job at explaining premiums and discounts on ETFs. There is a slight negative pattern, but it is not the most significant.

Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index v Real Premium

Discount

Change in Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index
Interbank Borrow Portion of the Cleveland Fed Financial Stress index does not provide information on ETF discounts because it appears to be randomly scattered.

*Interbank or Funding Markets Portion of Cleveland Financial Stress Index v Real Premium / Discount*
Interbank Funding Portion of the Cleveland Fed Stress index shows a pattern, but not one that can explain the direction of the premium or discount. It appears that as the funding portion is low, there are not big premiums or discounts. As the stress increases, however, there are more extreme premiums and discounts, but the direction appears unpredictable.

*St Louis Fed Stress Index v Real Premium / Discount*
As the St. Louis Fed Stress index is low, the ETF premiums and discounts tend to be small. As stress rises, the premiums and discounts become extreme, but they are not in any particular direction. This means that this measure may not be a good predictor of whether the funds are trading at premiums or discounts, but it may be useful to find the intensity of the ETF’s deviations from NAV.
Appendix C

Other Single Variable Explanations for Deviations

Appendix C includes all explanatory variables that attempt to predict the deviations from NAV in the ETF market. This appendix covers all the independent variables that were excluded from chapter 7.

Single Variable Regression Table Explaining “Real Deviation” Ordered By R Squared

<table>
<thead>
<tr>
<th>Explanatory Variables For &quot;Real Deviation&quot;</th>
<th>R Squared</th>
<th>Coefficient</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yesterday Deviation</td>
<td>0.7413329</td>
<td>0.89556636</td>
<td>0.000158</td>
</tr>
<tr>
<td>Dividend Yield</td>
<td>0.3030438</td>
<td>0.0017165</td>
<td>-0.00191</td>
</tr>
<tr>
<td>Financials Cost of Equity</td>
<td>0.2955414</td>
<td>0.0005649</td>
<td>-0.00549</td>
</tr>
<tr>
<td>ETF Average Volume</td>
<td>0.2688048</td>
<td>1.3459E-10</td>
<td>-0.00028</td>
</tr>
<tr>
<td>Barclays Junk Index</td>
<td>0.2626384</td>
<td>0.003511</td>
<td>-0.00047</td>
</tr>
<tr>
<td>St Louis Fed Stress Index</td>
<td>0.2620432</td>
<td>0.0008265</td>
<td>0.001752</td>
</tr>
<tr>
<td>Yesterday Discount</td>
<td>0.2517384</td>
<td>-0.5860017</td>
<td>0.00147</td>
</tr>
<tr>
<td>Average Daily Volume</td>
<td>0.2329664</td>
<td>5.0063E-10</td>
<td>-0.00028</td>
</tr>
<tr>
<td>Median P/E</td>
<td>0.2145246</td>
<td>-0.0004274</td>
<td>0.008826</td>
</tr>
<tr>
<td>VIX Index</td>
<td>0.1967801</td>
<td>9.3808E-05</td>
<td>-0.0037</td>
</tr>
<tr>
<td>Credit Suisse Fear Barometer</td>
<td>0.1777501</td>
<td>-0.0001594</td>
<td>0.005125</td>
</tr>
<tr>
<td>Interbank or Funding Markets Portion of Cleveland Financial Stress index</td>
<td>0.1750006</td>
<td>0.00030969</td>
<td>-0.0003</td>
</tr>
<tr>
<td>Conference Board Consumer Confidence Index</td>
<td>0.1578783</td>
<td>-3.659E-05</td>
<td>0.004125</td>
</tr>
<tr>
<td>Russell Vol</td>
<td>0.1492521</td>
<td>6.0231E-05</td>
<td>0.000109</td>
</tr>
<tr>
<td>Michigan Confidence</td>
<td>0.1460977</td>
<td>-7.74E-05</td>
<td>0.007474</td>
</tr>
<tr>
<td>S&amp;P Vol</td>
<td>0.1404326</td>
<td>6.6304E-05</td>
<td>0.00038</td>
</tr>
<tr>
<td>Market Crash Portion of Cleveland Fed Financial Stress Index</td>
<td>0.1309728</td>
<td>0.00011037</td>
<td>8.21E-05</td>
</tr>
<tr>
<td>Dow Vol</td>
<td>0.128706</td>
<td>7.0755E-05</td>
<td>0.000397</td>
</tr>
<tr>
<td>Cleveland Fed Financial Stress</td>
<td>0.1149263</td>
<td>0.00061389</td>
<td>0.001321</td>
</tr>
<tr>
<td>ABS: ETF Volume Change</td>
<td>0.110959</td>
<td>7.9366E-11</td>
<td></td>
</tr>
<tr>
<td>Bloomberg Confidence</td>
<td>0.1108908</td>
<td>-9.004E-05</td>
<td>0.004606</td>
</tr>
<tr>
<td>Financials Beta</td>
<td>0.1066827</td>
<td>0.00381887</td>
<td>-0.00319</td>
</tr>
<tr>
<td>Average WACC</td>
<td>0.0945671</td>
<td>0.00118533</td>
<td>-0.00949</td>
</tr>
<tr>
<td>Average Short Interest</td>
<td>0.0853329</td>
<td>-0.0009545</td>
<td>0.005148</td>
</tr>
<tr>
<td>ABS: Financials Beta Change</td>
<td>0.0754981</td>
<td>0.06318632</td>
<td></td>
</tr>
<tr>
<td>Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index</td>
<td>0.0647522</td>
<td>0.0005282</td>
<td>0.000802</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Ibox Investment Grade Index</td>
<td>0.0620016</td>
<td>-1.397E-05</td>
<td>0.004075</td>
</tr>
<tr>
<td>Average EPS</td>
<td>0.0513801</td>
<td>-0.000227</td>
<td>0.001821</td>
</tr>
<tr>
<td>Average P/B</td>
<td>0.0359507</td>
<td>-4.586E-05</td>
<td>0.001891</td>
</tr>
<tr>
<td>ABS: Russell Change%</td>
<td>0.0353989</td>
<td>0.04761455</td>
<td>0.001131</td>
</tr>
<tr>
<td>ABS: S&amp;P Change%</td>
<td>0.0350347</td>
<td>0.06320061</td>
<td>0.001153</td>
</tr>
<tr>
<td>Copper</td>
<td>0.0337356</td>
<td>-2.279E-07</td>
<td>0.003063</td>
</tr>
<tr>
<td>Dollar Strength</td>
<td>0.0334317</td>
<td>-8.677E-05</td>
<td>0.008575</td>
</tr>
<tr>
<td>Average Bid - Ask Spread %</td>
<td>0.0329984</td>
<td>0.0158296</td>
<td>0.001688</td>
</tr>
<tr>
<td>ABS: Dow Change%</td>
<td>0.0317656</td>
<td>0.06673267</td>
<td>0.001161</td>
</tr>
<tr>
<td>Average Debt/Equity</td>
<td>0.0312881</td>
<td>7.867E-06</td>
<td>0.000249</td>
</tr>
<tr>
<td>Oil</td>
<td>0.030907</td>
<td>-1.818E-05</td>
<td>0.002993</td>
</tr>
<tr>
<td>Liquidity Component to Cleveland Fed Financial Stress Index</td>
<td>0.0304012</td>
<td>0.004948</td>
<td>0.000149</td>
</tr>
<tr>
<td>Silver</td>
<td>0.0286834</td>
<td>-4.005E-05</td>
<td>0.002777</td>
</tr>
<tr>
<td>Financials WACC</td>
<td>0.0223119</td>
<td>-0.0002987</td>
<td>0.003485</td>
</tr>
<tr>
<td>Average P/E</td>
<td>0.0167362</td>
<td>-5.89E-06</td>
<td>0.001762</td>
</tr>
<tr>
<td>Fed Funds</td>
<td>0.0159442</td>
<td>-0.0001317</td>
<td>0.001719</td>
</tr>
<tr>
<td>Gold</td>
<td>0.0110423</td>
<td>-5.334E-05</td>
<td>0.002082</td>
</tr>
<tr>
<td>Retail ETF Ownership</td>
<td>0.0079073</td>
<td>9.926E-06</td>
<td>0.00024</td>
</tr>
<tr>
<td>Institutional ETF Ownership</td>
<td>0.0079073</td>
<td>-9.926E-06</td>
<td>0.001232</td>
</tr>
<tr>
<td>Insider Ownership</td>
<td>0.0076519</td>
<td>0.0058779</td>
<td>-0.00052</td>
</tr>
<tr>
<td>Change in Liquidity Cleveland Fed Financial Stress Index</td>
<td>0.0030195</td>
<td>0.02110819</td>
<td>0.001507</td>
</tr>
<tr>
<td>Financials WACC Change in %</td>
<td>0.0011751</td>
<td>0.00571486</td>
<td>0.001509</td>
</tr>
<tr>
<td>Institutional Ownership</td>
<td>0.0011751</td>
<td>-1.088E-05</td>
<td>0.001714</td>
</tr>
<tr>
<td>Change in Retail Ownership</td>
<td>0.0008853</td>
<td>-2.962E-05</td>
<td>0.000732</td>
</tr>
<tr>
<td>Change in Ibox Investment Grade Index %</td>
<td>0.0008332</td>
<td>0.00089387</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in Conference Board Consumer confidence index %</td>
<td>0.0007992</td>
<td>0.00248963</td>
<td>0.001509</td>
</tr>
<tr>
<td>% Retail owned (S and P)</td>
<td>0.0007941</td>
<td>9.7677E-06</td>
<td>0.000658</td>
</tr>
<tr>
<td>Change in Interbank or Funding Markets Portion of Cleveland Financial Stress index</td>
<td>0.0006704</td>
<td>-0.0051957</td>
<td>0.00151</td>
</tr>
<tr>
<td>Financial Cost of Debt</td>
<td>0.000634</td>
<td>6.0007E-05</td>
<td>0.00136</td>
</tr>
<tr>
<td>Change in Oil %</td>
<td>0.0005683</td>
<td>0.00278205</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in Dollar Strength %</td>
<td>0.0004944</td>
<td>-0.0118772</td>
<td>0.00151</td>
</tr>
<tr>
<td>Dow Vol Change in %</td>
<td>0.0004619</td>
<td>0.00420898</td>
<td>0.001506</td>
</tr>
<tr>
<td>Change in St Louis Fed Stress Index</td>
<td>0.0004604</td>
<td>4.619E-05</td>
<td>0.001511</td>
</tr>
<tr>
<td>Russell Vol Change in %</td>
<td>0.0004334</td>
<td>0.00461399</td>
<td>0.001506</td>
</tr>
<tr>
<td>Change in Copper %</td>
<td>0.0003685</td>
<td>0.00259046</td>
<td>0.001509</td>
</tr>
<tr>
<td>S&amp;P Vol Change in %</td>
<td>0.0003669</td>
<td>0.00421404</td>
<td>0.001506</td>
</tr>
<tr>
<td>Change in Retail ETF Ownership %</td>
<td>0.000315</td>
<td>-0.0010108</td>
<td>0.000732</td>
</tr>
<tr>
<td>ETF Volume Change</td>
<td>0.0002747</td>
<td>1.7981E-12</td>
<td>0.001509</td>
</tr>
<tr>
<td>Dividend Yield Change in %</td>
<td>0.0002636</td>
<td>-0.0011956</td>
<td>0.00151</td>
</tr>
<tr>
<td>Financials Cost of Equity Change in %</td>
<td>0.000248</td>
<td>0.00696458</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in Michigan Confidence %</td>
<td>0.0002303</td>
<td>0.00243811</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in Cleveland Fed Financial Stress</td>
<td>0.0002227</td>
<td>-7.07E-05</td>
<td>0.001512</td>
</tr>
<tr>
<td>Δ in Financials WACC</td>
<td>0.0001471</td>
<td>0.0022437</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in Barclays Junk Index %</td>
<td>0.0001302</td>
<td>0.00196051</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change in VIX Index %</td>
<td>0.000125</td>
<td>0.00150605</td>
<td>0.001507</td>
</tr>
<tr>
<td>Change in Cleveland Fed Financial Stress Index</td>
<td>5.804E-05</td>
<td>0.0052923</td>
<td>0.001508</td>
</tr>
<tr>
<td>Change in Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index</td>
<td>5.195E-05</td>
<td>-8.497E-05</td>
<td>0.001509</td>
</tr>
<tr>
<td>Financial Cost of Debt Change in %</td>
<td>2.362E-05</td>
<td>-3.12E-07</td>
<td>0.00151</td>
</tr>
<tr>
<td>Financial Cost of Debt Change %</td>
<td>2.312E-05</td>
<td>-0.0004728</td>
<td>0.0015</td>
</tr>
<tr>
<td>Average WACC Change %</td>
<td>1.855E-05</td>
<td>0.00846711</td>
<td>0.001509</td>
</tr>
<tr>
<td>Financial Cost of Equity Change %</td>
<td>1.562E-05</td>
<td>4.438E-05</td>
<td>0.00151</td>
</tr>
<tr>
<td>Dow Change %</td>
<td>1.562E-05</td>
<td>4.438E-05</td>
<td>0.00151</td>
</tr>
<tr>
<td>Change in Silver %</td>
<td>5.145E-06</td>
<td>-0.0001534</td>
<td>0.00151</td>
</tr>
<tr>
<td>Change in Bloomberg Confidence %</td>
<td>1.696E-06</td>
<td>8.1294E-05</td>
<td>0.001509</td>
</tr>
<tr>
<td>Change In Market Crash Portion of Cleveland Fed Financial Stress Index</td>
<td>1.081E-06</td>
<td>-1.207E-05</td>
<td>0.00151</td>
</tr>
<tr>
<td>Financial Cost of Debt Change %</td>
<td>9.706E-07</td>
<td>-0.0010211</td>
<td>0.00151</td>
</tr>
<tr>
<td>S&amp;P Change %</td>
<td>9.706E-07</td>
<td>-0.0010211</td>
<td>0.00151</td>
</tr>
<tr>
<td>Change in Credit Suisse Fear Barometer %</td>
<td>3.342E-07</td>
<td>-0.001473</td>
<td>0.00151</td>
</tr>
<tr>
<td>Financials Beta Change %</td>
<td>7.202E-08</td>
<td>-0.0006556</td>
<td>0.00151</td>
</tr>
<tr>
<td>Russell Change %</td>
<td>7.202E-08</td>
<td>-0.0006556</td>
<td>0.00151</td>
</tr>
</tbody>
</table>

Explanatory Variables for “Real Premiums / Discounts” and Real Deviations ordered as observed:

*Average Bid/Ask Spread % [of the Stocks listed on the S and P]*
*Average WACC*
*Average D/E*
*Average Short Interest*
*Median P/E*
*Average P/E*
Average P/B
Average Daily Volume
Financials WACC
Financial Cost of Debt
Financials Cost of Equity
Financials Beta
S&P Vol
Dow Vol
Russell Vol
Institutional Ownership
Insider Ownership
Dividend Yield
Average EPS
S&P, Dow, and Russell Change%
VIX Index
Credit Suisse Fear Barometer
Fed Funds
Dollar Strength
Barclays Junk Index
Ibox Investment Grade Index
Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer Confidence Index
Gold, Silver Copper, and Oil
ETF Average Volume
Liquidity Component to Cleveland Fed Financial Stress Index
Cleveland Fed Financial Stress
Market Crash Portion of Cleveland Fed Financial Stress Index
Interbank Borrow Rate Portion of Cleveland Fed Financial Stress Index
Interbank or Funding Markets Portion of Cleveland Financial Stress index
St Louis Fed Stress Index

Explaining Deviations

Yesterday Deviation v Real Deviation
Predictably, yesterday deviation from NAV is a good predictor of what the next day’s deviation will be. It appears to be a relatively tight linear regression so no transformations will be needed.

*Average Bid/Ask Spread % [of the Stocks listed on the S and P] v Real Deviation*
Bid-ask spread and change in bid-ask spread seem to produce random scatter and are not a good explanation for real deviations. This will not be examined further.

*Average WACC v Real Deviation*
Average WACC seems to have a relatively weak positive relationship with real deviations of ETFs from NAV. The only problem with the distribution is that the deviations seems to fan out after the WACC increases past 8.5. Increased WACC was predicted to cause larger deviations because it makes it more difficult for institutional investors to finance arbitrage operations.

The change in WACC does not seem to be a good predictor of the real deviations and it will not be examined further.

Average D/E v Real Deviation
Both average debt to equity and debt to equity changes seem to be distributed randomly and do not make for good predictor of deviations from NAV.

*Average Short Interest v Real Deviation*
There is a linear relationship between short interest and deviations from NAV, but it is the opposite of what was expected. Hypothetically, short interest should have acted as a proxy for investor panic. As more investors are worried, short interest should rise. This is not the case with this data set. An alternative hypothesis could be that if more people are shorting stocks, then traders are bringing the traded price of ETFs closer to their NAV.
The change in short interest shows random scatter.

*Median P/E v Real Deviation*
Median P/E was predicted to be a proxy for investor sentiment. As it rises, investors are more confident in firms prospects for future earnings. When investors are more confident, ETFs end to trade closer to their NAV, so as P/E rises, deviations from NAV falls.

On the change in median P/E, a linear regression is not appropriate because the pattern shows a V. shape. Instead, a second order polynomial regression looks like it will work. From the graph, it appears that, regardless of direction, as P/E changes grow more extreme, the deviations from NAV in the ETF market increases.

*Average P/E v Real Deviation*
Average P/E average P/E changes are skewed by extreme numbers and are not appropriate predictors of ETF deviations from NAV.

*Average P/B v Real Deviation*
Average P/B changes are skewed by extreme numbers and are not appropriate predictors of ETF deviations from NAV.

Average Daily Volume v Real Deviation
The average daily volume of the S&P has a positive linear relationship with ETF deviations from NAV. The relationship is one of the stronger ones and warrants further examination. The rational for the relationship is clear. As more investors are trading securities, there is an increased chance of mispricing as volume can overwhelm institutional investors’ abilities to take advantage of arbitrage in the marketplace.

Change in volume shows no clear relationship and will not be considered further.

*Financials WACC v Real Deviation*
Financials WACC shows a weak linear relationship with ETF deviations from NAV. Unfortunately, the direction of the relationship is not what was predicted. As financials WACC increases, the graph shows deviations decreasing. Logically, the deviations should increase because it would be more expensive for banks to fund divisions to take advantage of arbitrage in the ETF markets.
The financials WACC change does not provide valuable information because it does not change that often.

**Financial Cost of Debt v Real Deviation**

![Graph showing Financial Cost of Debt v Real Deviation](image_url)

- Financial Cost of Debt
- Real Deviation
- $R^2 = 0.0006$
There is no clear relationship between financials cost of debt and deviations from NAV.

Financials Cost of Equity v Real Deviation
Financials cost of equity seems to be a strong indicator of real deviations from NAV. It was originally predicted that as cost of equity increases banks would be less willing to finance ETF operations. This prediction is true according to the graph. As cost of equity increases, deviations also increase.

This is not true with the financials cost of equity change. Changes are not recorded often, so the graph does not produce an accurate picture of how financials cost of equity changes affect ETF divergence.

Financials Beta v Real Deviation
Logically, as beta rises for financials, the market is becoming more unstable, so deviations should increase. Financials beta shows a slight positive relationship with ETF divergence from NAV. The data verifies this assumption. Financials beta change shows no linear relationship with ETF divergence.
A polynomial regression, however, shows a very interesting relationship. Regardless of direction of the beta change, as changes in beta gets more extreme the ETF deviations gets more pronounced.

*S&P Vol v Real Deviation*
S&P volatility shows a clear positive linear relationship with real deviations from NAV. This is expected because as volatility increases it may be impossible for banks to execute arbitrage transactions on the ETF market because prices may move before the transaction is complete.

S&P volatility change shows a relatively weak correlation with deviations from NAV. It will not be examined further.

*Dow Vol v Real Deviation*
Dow volatility tells the same story as S&P volatility. As volatility increases, deviations from NAV also tend to increase.

*Russell Vol v Real Deviation*
Russell volatility tells the same story as S&P volatility. As volatility increases, deviations from NAV also tend to increase.

*Institutional Ownership v Real Deviation*
Prior research from closed end funds implied that increased institutional ownership of the S&P/Russell led to lower deviations from NAV in CEFs. The CEF institutional ownership trends do not seem to carry to ETFs. There is not a strong relationship between institutional ownership and deviations for ETFs.

*Insider Ownership v Real Deviation*
Like institutional ownership, insider ownership shows very little relationship with deviations from NAV.

*Dividend Yield v Real Deviation*
If dividend yield is a proxy from market pessimism, the hypothesis that more pessimism breeds large deviations holds true. There appears to be a moderate positive correlation between average S&P dividend yield and ETF deviations. As yield increase, deviations also increase.
The same is not true for the dividend yield change. Yield changes offer no insight into the real deviations.

*Average EPS v Real Deviation*

As with most averages, EPS is skewed by extreme numbers, so it provides no predictive information to what the real deviations from NAV will be.

*S&P, Dow, and Russell Change% v Real Deviation*
For all of the index change charts, the pattern is roughly the same. As the index changes extremely up or extremely down, the deviations from NAV increases. This is expected because violent moves are expected to overwhelm institutional investors who push the ETFs to track the underlining indexes.
As the VIX index increases, the deviations from NAV also increase. This is expected as VIX is called the fear index. As investors are more fearful there are expected to be less market participants who will force the ETFs to trade efficiently close to their NAV.
The Credit Suisse Fear Barometer is similar to the VIX in that it is generated using option prices to find volatility. Unlike the VIX, however, it only used the downside risk pricing. This creates an odd system that turns out to have an opposite relationship with the deviations from NAV. As fear increases from the Credit Suisse Fear Barometer, deviations appear to fall. This is
highly abnormal and runs very counter to idea that fear breeds deviation. It is important to note that the Credit Suisse Fear Barometer is supposed to be forward looking (to track the fear of the next few months) rather than the fear on spot (like VIX).

The change in the Fear Barometer seems to produce random scatter.

*Fed Funds v Real Deviation*
While there is a pattern in Fed Funds Rate and the deviation, it does not look like it is either linear or explanatory.

*Dollar Strength v Real Deviation*
The dollar strength provides no predictive power to the ETF deviations from NAV.

*Barclays Junk Index v Real Deviation*
It appears that the Barclays Junk index has a relatively strong linear positive correlation with the ETF deviations from NAV. As the Junk index performs better, the ETFs seem to deviate from NAV more. This is counter to what is predicted in theory. Theoretically, as junk does better, deviations should decrease because investors should be more confident in markets which is expected to drive deviations down.

*Ibox Investment Grade Index v Real Deviation*
Also surprising is the behavior of investment grade securities. As the Ibox Investment grade index performs better, the deviations from NAV tend to decrease. This is surprising because one would believe that investment grade securities perform better in times of crisis when deviation should be higher.
Michigan Confidence, Bloomberg Confidence, and Conference Board Consumer Confidence Index vs Real Deviation

**Michigan Confidence**

\[ R^2 = 0.1494 \]

**Bloomberg Confidence**

\[ R^2 = 0.1112 \]
All of the measures of confidence selected for this exercise to explain ETF deviations from NAV fail to produce meaningful results with the exception of the lower end of the Michigan Confidence index.

*Gold, Silver Copper, and Oil v Real Deviation*
Silver

Change in Silver %
All commodity types produce no real pattern that can be used to interpret the deviations from NAV. The only recognizable pattern is one in the gold deviations chart, but it appears to be not useful because the pattern reverses at the 600$/oz mark.
ETF Average Volume v Real Deviation

**ETF Average Volume**

- $R^2 = 0.2889$

**ETF Volume Change**

- $R^2 = 0.1502$
- $R^2 = 0.0001$
ETF volume looks to be one of the strongest positive linear correlations to NAV deviation. As ETF volume increases, the deviations increase as well. This is consistent with the hypothesis that increased volume overwhelms institutional investor’s ability to correct for ETF mispricing.

Additionally, the change in ETF volume also seems to be significant. As ETF volume changes (regardless of direction), the deviations from NAV also increase. Again, this is consistent with the theory that overwhelming volume can enable the deviations from NAV.

*Liquidity Component to Cleveland Fed Financial Stress Index v Real Deviation*
The Liquidity Portion of the Cleveland Fed Financial Stress index seems to produce random scatter and will not be examined further.

_Cleveland Fed Financial Stress v Real Deviation_
The Cleveland Fed Financial Stress index has a moderate positive linear correlation with the average ETF deviations from NAV. This is consistent with the hypothesis that as financial stress increase, deviations tend to be larger due to tying up institutional investors in other activities during stress periods.

*Market Crash Portion of Cleveland Fed Financial Stress Index v Real Deviation*
The Market Crash Portion of the Cleveland Fed Financial Stress index shows a moderate to strong positive correlation with NAV. This is consistent with the hypothesis that market crashes and general financial stress causes deviations from NAV. This could be because during market crashes banks are unable to participate in the ETF market to correct for errors.

The change in this measure, however produces random scatter.
The Interbank Borrow Portion of the Cleveland Fed Financial Stress Index shows that for most values under 2.5 the interbank rate does not do a good job at predicting the ETF deviations from...
NAV. For values greater than 2.5, however, there is a relatively strong positive linear correlation. This indicates that as the interbank stress increases there are larger deviations from NAV in the ETF market. This is not surprising as financial stress could be one of the things that inhibit market participants from correcting mispricing in the ETF market.

Change in the Interbank Borrow rate, however is unpredictive.

*Interbank or Funding Markets Portion of Cleveland Financial Stress index v Real Deviation*

---

![](image)
The Interbank Funding Rate Portion of the Cleveland Fed Financial Stress index provides an extraordinarily strong positive linear correlation with the deviations from NAV on the ETF market. This is consistent with the idea that additional financial stress inhibits banks from acting to correct the ETF market.

*St Louis Fed Stress Index v Real Deviation*
The St Louis Stress index provides a very strong linear relationship with ETF deviations from NAV. This is consistent with the idea that financial stress breeds larger deviations from NAV due to inhibiting bank’s abilities to correct pricing errors in the ETF market.
BIBLIOGRAPHY


Education
Major(s) and Minor(s): Bachelors of Science in Finance & Economics
Honors: Finance

Thesis Title: ETF DIVERGANCE FROM NET ASSET VALUE
Thesis Supervisor: David Haushalter

Work Experience:

Summer 2015
Position- Summer Analyst
Company- Goldman Sachs
Supervisor’s Name- Lindsey Morfin

Summer 2014
Position- Corporate Strategy & Business Development Intern
Company- AmerisourceBergen
Supervisor’s Name- Larry Marsh