SECTOR EXCHANGE TRADED FUNDS: AN ANALYSIS OF FUND FLOW AND RETURN

Abdullah Noman, Nicholls State University
Shari Lawrence, Nicholls State University
John Lajaunie, Nicholls State University
Michelle Connell, University of Texas at Dallas

ABSTRACT

The relationship between ETF fund flow and return, as well as the relevance of this relationship for positive and negative return is analyzed in this paper. Additionally, the duration of these relationships is considered. Two hypotheses are developed and tested to determine if past return predicts future fund flow of ETFs and whether ETF fund flow predicts future return. The results indicate that past return does have a significant effect on future fund flow to and from ETFs. However, the results overall do not indicate a clear causal pattern. With regard to the duration of the relationship, increased fund flow generally leads to significantly higher returns in the next month, but tapers off and is not significant in the second and third months following the event. Finally, negative fund flow has stronger predictive ability compared to positive flow, indicating that fund outflow following poor performance is greater than fund inflow following good performance. This implies that investors are more likely to engage in relatively significant selling in the wake of negative return than aggressive buying to "chase" positive return. This finding is consistent with the assumption of investor risk aversion. **JEL Classification:** C33, G11, G12

INTRODUCTION

The purpose of this paper is to evaluate the relationship between fund flow and return for Exchange Traded Funds (ETFs). Two hypotheses are developed and tested. The first hypothesis considers the relationship between fund return and fund flow of ETFs. In other words, does past return predict future fund flow of ETFs? The second hypothesis considers the relationship between past ETF fund flow and future return. If the results fail to reject the second hypothesis, then the significance of the relationship for negative and positive future return is tested. Finally, if the existence of this relationship for both types of return, positive and negative, cannot be rejected, then a test for the duration is conducted. In other words, does the "smart money" investor get into the market via ETFs first, followed by the "dumb money" investor?

The aforementioned questions are considered relevant because of the increasing use of ETFs in today's volatile market. Until recently, individual and institutional investors relied primarily on mutual funds or separately managed accounts (e.g. funds customized for each investor). ETFs have become increasingly popular in recent years due to cost efficiencies and a general tendency towards a more passive management strategy by investors. As a result, money has flowed from the mutual fund industry to ETFs. It is expected that this trend will continue into the future with some investment professionals projecting that ETF NAV will eventually exceed mutual fund NAV. Some analysts estimate the future ETF NAV to reach $15 trillion by 2024 (Hougan and Nadig, 2014). Thus, research focusing on the fund flow into ETFs and the motive for fund flow is an important endeavor.

Initially, ETFs allowed the investor/buyer to participate in the ownership of market indices such as the S&P 500 or the NASDAQ. Over time, ETFs have become more inclusive and specific; ETFs can now be purchased to mirror a variety of industry sectors, fixed income securities, commodities and leveraged strategies. In this paper, a data set was constructed that represents the sector ETFs. The sample period is from January 2005 to December 2013 with a monthly frequency. The selection of the sample period was based primarily on data availability and completeness for the sector ETFs managed by SPDR, iShares and Vanguard.

The following sections of the paper include a brief literature review and discussion of research questions, data and methodology, analysis of results, and a conclusion.

BRIEF LITERATURE REVIEW AND RESEARCH QUESTIONS

The first ETF was introduced in 1990. It was designed to track the Toronto Stock Exchange (TSE-35) stock index. Prior to this however, the concept of having a basket of securities that investors could trade effectively and efficiently was being analyzed. After some trial and error, the American Stock Exchange introduced the Standard and Poor's Depository Receipts (SPDR) in 1993. Its purpose was to provide investors with a cost effective way of investing in a basket that tracks the S&P 500 index. This ETF proved to be so popular that it had grown to a Total Net Assets of $2.76 trillion by the end of November 2014 (Zacks Funds, 2014). The popularity of SPDRs has led to the development of numerous basket type products, providing investors a wide variety of investment choices at a lower cost relative to most mutual funds.

Initially, ETF products were based on market wide equity indices. This was soon expanded to include fixed income indices and sector indices. As the market for these ETF products became saturated, products based on other classes such as commodities and currencies were introduced. Recently, the development of actively managed ETF products, as well as leveraged and long-short ETFs have been introduced to the market. Currently, there are over 1,659 ETFs an investor can choose from (Zacks Funds, 2014).

Given the growth in assets, ETFs appear to have become the investment vehicle of choice compared to mutual funds. ETFs have several benefits relative to costs, liquidity, flexibility and tax management when compared to mutual funds. Typically, ETF fees range from 0.10% to 1.25%. In comparison, mutual fund fees can range from .01% to 10%. (Pareto, 2015). The importance of lower fees has been a frequent topic by investment professionals. John Bogle, of Vanguard funds, has stated that fees erase a large piece of the compounding effect of a portfolio. For example, if a 2% fee is charged, an investor can potentially lose more than 60% of a portfolio's value over a fifty-year time horizon (Heinzel, 2013). In addition, actively managed mutual funds frequently underperform their benchmark indices. For example 80-90% of the actively managed funds did not beat their respect benchmark index in 2014 (Constable, N. and Kadnar, M., 2015). Given these issues, there has been a definitive shift towards passive management as the norm in the investment world, and the development of ETFs has facilitated this shift.

Regarding flexibility and liquidity benefits, ETFs, unlike mutual funds, can be traded like individual equities. In other words, investors can freely buy and sell ETFs throughout the day, whereas mutual fund investors receive the end of the day price. Another benefit of ETFs includes lower taxes. Whereas mutual fund owners are "billed" for the capital gains within the fund as the individual holdings are sold, ETF holders are not taxed until redemption.

There are some disadvantages to using ETFs, however. One is that customers must pay a commission, in most cases, when buying or selling. Another is the bid-ask spread, which can be significant for ETFs with low liquidity. As previously stated, most ETFs have relatively low fees; however, not all ETF fees are low. Some actively managed ETF fees can be somewhat higher, ranging from 1-2 percent. Finally, leveraged ETFs can experience something called "decay" in which return is adversely affected (due to the leverage component) if the funds are held longer than just a few days. However, the disadvantages of ETFs do not appear to outweigh the advantages as evidenced by their increasing popularity with investors.

Most of the research on ETFs to date has focused on three areas: price efficiency, tracking ability and performance, and the effects on the underlying securities (Charupat and Miu, 2013). Regarding price efficiency, studies indicate that price deviations are very small on average and the size of such deviations is generally related to the underlying NAV's (Engle and Sarkar, 2006). Tracking error is defined as the difference between the return of the ETF and the corresponding return on its underlying benchmark index. There have been numerous studies on tracking errors and performance (Agapova, 2010; Elton, Gruber, Comer and Li, 2002; Kostovetsky, 2003). The consensus is, in general, that tracking error has a small impact on ETF performance, on average. Regarding any effects on underlying securities, Gorton and Pennacchi (1993) have studied this area extensively. The belief is that there will be a migration out of individual equities and into ETFs because of the reduction of "firm specific" risk in ETFs. This may result in individual securities becoming less liquid and having increased bid-ask spreads. Thus far, however, the findings regarding the effects of ETFs on underlying securities have been inconclusive.

While very little research focuses specifically on ETF fund flow, there have been numerous studies comparing ETFs and mutual funds (Elton, Gruber, and Busse, 2004; Friesen and Sapp, 2007; Berk and Green, 2004; Frazzini and Lamont, 2008). The conclusions of these studies are all in general agreement that the cash flow behavior in mutual funds is very similar to that of ETFs. Recently, Clifford and Fulkerson and Jordan (2014) offered some insight into the flow of funds to ETFs using panel data. In addition, evidence suggests that both ETF and mutual fund flow demonstrates "return chasing" behavior, implying that significant positive return are often followed by significant fund inflow.

In this paper, two primary research questions are addressed: first, whether past return predicts future fund flow and second, whether past fund flow predicts future return. Based on these questions, the empirical specification to examine the relationship between fund flow and return is developed. Regarding the question of past return predicting future flow, in an efficient market past return should not predict future fund flow, as rational investors do not base their investment decisions on past performance. Conversely, rational investors tend to focus on expected future performance. In prior empirical work, however, it is reported that there is indeed a relationship between return and flow, which can have either rational or behavioral explanations (e.g. Sirri and Tufano, 1998; Del Guercio and Tkac, 2002).

Under the two behavioral paradigms considered in this study, past return may positively predict future flow, as investors may suffer from psychosocial biases such as representative heuristics (Tversky and Kahneman, 1974). An alternative explanation for the significant return–flow relationship is based on rational learning (Berk and Green, 2004, Bollen, 2007). The idea of rational learning hinges on investors' gradual learning of past investment outcomes, causing autocorrelation in the innovations of the data generating process. Empirically, this translates into predictability of future flow based upon past return. In the event, a significant relationship between flow and return is found in our sample. Hence, there is the question of distinguishing between the behavioral versus rational explanation. One possible solution is to use the asymmetric response model that would allow for the examination of flow response to positive and negative past returns. If flow responds differently to past positive and negative returns, then this would suggest the behavioral explanation is more applicable. Conversely, if there is no difference, then the results would suggest the rational explanation is more appropriate.

The second question in this line of research is whether investors can make the rational choice when they invest in an ETF. This can be examined by analyzing the flow of funds to the ETF. The question is whether flow of funds to the ETF predicts the future return. If flow of funds is not significant, then it does not predict future return. As a result, past fund flow has no information regarding future return. On the other hand, if fund flow is found to predict future return, then two different explanations are plausible. First, if past fund flow positively predicts future return, then there is support for a "smart money" investor theory. Investors are able to chase return by allocating their money to winning ETFs. However, when past fund flow negatively predicts future return, then a "dumb money" investor theory is plausible.

DATA AND METHODOLOGY

Monthly data were obtained for three sector ETFs managed by SPDR, iShares and Vanguard fund families from the Center for Research in Security Prices (CRSP) database. The list of ETFs used in this study is provided in Table 1 of the appendix. The sample period spanned January 2005 to December 2013. Since different fund families began their ETFs at different times, a sample period was selected in which data was available for all funds. The data included adjusted closing price, volume, net asset value, outstanding shares, total net assets and premium/discount. The model specifications were estimated for each fund family by pooling data for each time series across all ETFs within each family. This created three subsamples: SPDR, iShares and Vanguard. On average, SPDR funds have more assets under management relative to iShares and Vanguard funds. In addition, a full sample estimation was conducted that included all 29 ETFs. This allows for the results in our study to vary across fund families, while also obtaining the overall picture of the sector ETFs as a whole.

Panel Vector Autoregression (PVAR)

A Panel Vector Autoregression (PVAR) was specified to study the relationship between fund flow and return. Another variable included in the VAR is "premium", which measures the difference between the price of the ETF and value of the underlying asset. Premium has the potential to capture a number of market and investor related characteristics that simultaneously affect flow and return, and therefore is included in the analysis (Delcoure and Zhong, 2007).

Consider a vector of three potentially endogenous variables $Z_{it}=[Flow_{it}, Ret_{it}, Prem_{it}]$, where, $Flow_{it+1}=\{TNA_{it+1}- [TNA_{it}\times(1+R_{it+1})]\}/TNA_{it}$, is the percentage net flow of funds to an ETF, is its total net asset and Ret_{it} is return over the previous period as reported in CRSP, is the premium (or discount) of a fund's price over its net asset value (NAV). The unrestricted VAR in the level with these variables can be written as:

$$Z_{it}=A_0+A_1\, Z_{(it-1)}+f_i+U_{it} \qquad (1)$$

where, and are vectors of constant and slope coefficients, and f is individual ETF–specific fixed effects, and other variables are defined as before. The vector of error terms, , are allowed to have unrestricted interaction among them. Panel VAR with individual fixed effects, however, would introduce bias in slope estimates. This bias is a result of the demeaning procedure in the fixed effects method (Arellano and Bover, 1995). To correct for this bias, we use the 'Helmert Transformation' following Love and Zicchino (2006). Essentially, this method implements forward demeaning of the variables instead of their regular demeaning, as done in fixed effects estimation. In order to explore the existence of any asymmetric relationship among the variables, a specification is used which is outlined in Exhibit 1 at the end of the appendix.

RESULTS AND ANALYSIS

The descriptive statistics are presented in Table 2 of the appendix. The first and second columns illustrate monthly average net flow and standard deviation. The SPDR funds have higher net average inflow for the study period, followed by Vanguard and iShares. In addition, SPDR funds have the highest volatility. The second and third columns illustrate average monthly return, all of which are positive for the total time period studied. Regarding premium, in the fifth column, the results indicate that the majority of SPDR funds are selling at a discount, with ishares and Vanguard to a lesser extent. The remaining columns present fund size, expense ratio, turnover, and bid–ask spread. The SPDR funds have a lower bid–ask spread compared to iShares and Vanguard, indicating higher liquidity.

The results of the panel VAR are reported in Table 3 of the appendix. As previously stated, the objective of this paper is to determine the relationship between return and flow given the specific sample data. The first question addressed is regarding past return predicting future flow. In particular, for the equation, the coefficients of the variable, , where the subscript *j* represents number of lags, were analyzed. In the full sample, the second and the third lags are significant. The coefficients are 0.688 (*t*–statistic: 3.90) and 0.694 (*t*–statistics 2.83), respectively. For the SPDR subsample, there is no evidence of past return affecting future flow at any lag, as all coefficients are insignificant. For iShares, the first and third lags are significant, while for Vanguard, second and third lags are significant. Overall, the results indicate that past return does indeed affect the future flow of funds to ETFs. However, the results did not provide a consistent pattern of positive return leading to increased flow and negative return leading to decreased flow.

The second question is whether past flow predicts future return. The focus of this question is the coefficients for with respect to the equation. For the full sample, the coefficients are significant at all three lags. The results are positive for the first lag (0.010, *t*–statistic 2.84), but negative for both the second (–0.005, *t*–statistic –2.49) and third lags (–0.010, *t*–statistics –2.92). Similar results are obtained for the Vanguard subsample. For the SPDR subsample, coefficients are significant in lags 2 and 3, with negative signs, while no significant coefficients are found for the iShares subsample. The overall results indicate that increased fund flow generally predicts higher return in the next month, but eventually lower return for the next 2 to 3 months. A possible explanation of the results is that flow is "smart" in the short horizon (1 month) but "dumb" over a comparatively longer horizon (in the second and third months). An interesting aspect of this result is the increasing magnitude of negative coefficients with each lag. In other words, the advantage of return chasing diminishes over time. Finally, the results also indicate that the opposite case is true, where decreased flow leads to lower return in the first month with the magnitude diminishing in the second and third months. Turning to the third variable (premium) in the VAR system, the results on how premium influences, and in turn is influenced by, the other two variables are briefly presented here. While premium is largely unaffected by the other two variables, it does have a significant impact on both flow and return. This finding indicates suitability of premium in the VAR specification above.

Variance Decomposition

Variance decomposition of the PVAR can help us gain further insight into the results of Table 3. In particular, the percent of variance in a response variable accumulated over 10 periods (i.e. months) resulting from shocks to impulse variables are presented in Table 4of the appendix. The variance decomposition results can be summarized as follows. In the Full sample, the variation in flow is almost entirely caused by shocks to itself, while slightly affected by return, but not by premium. On the other hand, variation in return is mostly due to shocks to itself (86.34%) and but also affected modestly by shocks in premium (11.81%). Finally, variation in premium is entirely explained by shocks to itself. For SPDR and iShares, variation in flow of funds is almost entirely (over 99%) explained by themselves. This is also true of the full sample. Shocks to return, or to premium, have little impact on variation in flow. As for the Vanguard sample, variance in return has a sizable impact (11.15%) on flow.

The variance decomposition analysis is done by maintaining the same order of variables as in the original panel VAR specification. Since the ordering of variables can have implications for the results in this section, the variance decomposition analysis was rerun with differing order of the variables. While not presented here, the results are found to be very similar.

Granger Causality Tests

Next, Granger causality tests were conducted in order to ensure further robustness of our analysis. The null hypothesis is that a set of excluded variables do not Granger-cause the left hand side variable. On the other hand, an alternative hypothesis is that the right hand side variables do Granger-cause the left hand side variable. In addition, these tests allow for tests of significance of the panel VAR model by excluding all right hand side variables for an equation. The panel VAR specification in equation (1) is used for the Granger causality tests. Within this framework, each of the potentially endogenous variables in the system is tested. The results are presented in Table 5 of the appendix

As in the full sample, one can see that the null hypotheses have been rejected for all cases, indicating the existence of significant relationships among the variables in the system. The results reconfirm the findings reported in the previous tables. Given the empirical evidence of bidirectional causality, a VAR specification is the appropriate choice.

Granger causality results can also be used to verify model significance for each of the three equations in the VAR system. For this purpose, we look at the row labeled "All" which presents the Wald statistics under the joint null hypothesis, where all explanatory variables are insignificant. A low p – value will reject the null hypothesis and provide evidence in favor of the VAR specification. First, we consider the equation for which the null is rejected for the full sample as well as all subsamples except the SPDR subsample. This indicates goodness of the particular specification. Second, for the equation, we are able to reject the null in all cases and obtain further support for the VAR specification. Finally, for the equation, the results are inconclusive, as the null is rejected for iShares and Vanguard subsamples, while not rejected for SPDR subsample and the full sample. Overall, the VAR specification is supported from these Wald statistics. While the results suggest that premium affected both flow and return in the system, it does not appear to be significantly influenced by them. It can, therefore, be argued that premium is more exogenous than the other two variables in the system. As a result, it was placed at the end in the variance decomposition analysis.

Asymmetric Response Model

The significant empirical relationship between return and funds flow may be due to either rational reasons or behavioral reasons. In order to distinguish between the two explanations for the return–fund flow relationship, an asymmetric response model was specified in which positive and negative values of the right hand side variables were analyzed to determine the structure of the effect. In other words, do the results clearly indicate whether there is any asymmetric relationship between return and fund flow with respect to the positive and negative values? A detailed description of the asymmetric response model is outlined in Exhibit 1 of appendix.

The results of the asymmetric response model are presented in Table 6 of the appendix. Regarding the question of whether positive or negative return predicts future flow differently, the coefficients of interest are the and variables in the equation.

The coefficients were insignificant in most cases with the exception of the Vanguard subsample. Thus, positive and negative past return appears to be affecting fund flow in the same manner for most subsamples. Combined with the results in the previous table, it would seem plausible that past return does indeed influence future fund flow. It is also likely this result can be attributed to a rational learning model such as those mentioned in Berk and Green, (2004) and Bollen, (2007), rather than the behavioral explanations.

Next, to see if positive and negative funds flow predicts return in an asymmetric manner, the coefficients on the and variables for the equation were analyzed. The coefficients were positive and significant for the variable for the full sample as well as for all subsamples. This suggests that positive flow appears to have a significant impact on future return while negative flow does not. However, this influence is limited to the first lag only. Combined with the evidence of a "smart money" effect reported in Table 3, this finding further supports the idea of dynamic loss aversion (O'Connell and Teo, 2009) whereby the funds flow is negative and significant following poor performance. This result can be explained using the behavioral biases introduced in (Kahneman and Tversky, 1979).

CONCLUDING REMARKS

The objective of this study is to examine the relationship between flow of funds and return, and vice versa, in a sample of sector exchange traded funds (ETFs). Specifically, this paper addresses two primary research questions. The first is whether past return predicts future fund flow. The second is whether flow of funds to the ETF predicts the future return. Answers to these questions are analyzed from the perspective of behavioral versus rational theories of investment with special reference to the smart money literature.

Monthly data are used to form three subsamples and one full sample consisting of 31 sector ETFs. A system of vector autoregression (VAR) is specified in three endogenous variables, namely flow of funds, return, and premium, where the third variable is considered to have an impact on the other two. In order to increase the power of the tests, a panel estimation method has been utilized for each subsample and for the full sample.

The findings of the study can be summarized as follows. First, significant bidirectional causality exists regarding return and fund flow in our sample period. Specifically, there is significant evidence of past return predicting the future fund flow. Additional analysis reveals that this relationship is based on a rational investor learning process. Second, significant evidence of fund flow chasing future return is also reported. As for the theory of smart money investors chasing return, the results indicate that there is some opportunity in the short run, but with time this fund flow becomes dominated by the "dumb" money investors, as return decreases.

This paper has a couple of limitations that also point to the direction of future research. First, the sample is limited to three fund families that issue and manage sector ETFs. However, there are many other sector ETFs that are popular. Future research could extend the data set to include all sector ETFs to examine whether or not the results from this analysis hold for a more comprehensive sample. Second, it would be interesting to see the relationship among return, flow, and premium during periods of financial crisis and high volatility. Any future endeavors would add more insight to the existing literature if such analysis is done with a more comprehensive sample as well as during periods of market volatility.

REFERNCES

Agapova, A. (2010). Conventional Mutual Index Funds Versus Exchange-Traded Funds. *Journal of Financial Markets*, *14*(2), 323-343.

Arellano, M., and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics*, *68*(1), 29-51.

Berk, J. B. and Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, *112* (6), 1269 – 1295.

Bollen, N. P. (2007). Mutual Fund Attributes and Investor Behavior. *Journal of Financial and Quantitative Analysis*, *42*(03), 683-708.

Charupat, N. and Miu, P. (2013). Recent Developments in Exchange-Traded Fund Literature: Pricing Efficiency, Tracking Ability, and Effects on Underlying Securities. *Managerial Finance*, *39*, 427-443.

Clifford, C., Fulkerson, J., and Jordan, B. (2014). What Drives ETF Flow? *Financial Review,* 49(3), 619–642.

Constable, N. and Kadnar, M. (2015). GMO White Paper: Is Skill Dead. Retrieved 3/1/15 from http://news.easybranches.com/gmo-white-paper-is-skill-dead/

Del Guercio, D., & Tkac, P. A. (2002). The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds vs. Pension Funds. *Journal of Financial and Quantitative Analysis*, 37(04), 523-557.

Delcoure, N. and Zhong, M. (2007). On the Premiums of iShares. *Journal of Empirical Finance*, *14*(2), 168-195.

Elton, E. J., Gruber, M. J. and Busse, J. A. (2004). Are Investors Rational? Choices Among Index Funds. *Journal of Finance*, 59(1), 261-288.

Elton, E. J., Gruber, M. J., Comer, G. and Li, K. (2002). Spiders: Where Are the Bugs? *The Journal of Business*, 75(3), 453-472.

Engle, R., & Sarkar, D. (2006). Premiums-Discounts and Exchange Traded Funds. *Journal of Derivatives*, 13(4), 27.

Frazzini, A., & Lamont, O. A. (2008). Dumb Money: Mutual Fund Flow and the Cross-Section of Stock Return. *Journal of Financial Economics*, *88*(2), 299-322.

Friesen, G. C., and Sapp, T. R. (2007). Mutual Fund Flows and Investor Returns: An Empirical Examination of Fund Investor Timing Ability. *Journal of Banking & Finance*, *31*(9), 2796-2816.

Gorton, G. B., & Pennacchi, G. G. (1993). Security Baskets and Index-Linked Securities. *Journal of Business*, 1-27.

Hjalmarsson, E. (2010). Predicting Global Stock Return. *Journal of Financial and Quantitative Analysis*, 45(01), 49-80.

Hougan, M. and Nadig, D. (2014) ETFs' Future: Huge Growth. Retrieved 3/4/15 from These questions were asked because of the current dominance of ETFs. Retrieved 3/4/15 from http://www.etf.com/sections/features/21150-the-future-of-etfs-why-we-expect-huge-growth.html?nopaging=1

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kostovetsky, L. (2003). Index Mutual Funds and Exchange-Traded Funds. *Journal of Portfolio Management*, *29*(4), 80-92.

Love, I., and Zicchino, L. (2006). Financial Development and Dynamic Investment Behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, *46*(2), 190-210.

O'Connell, P. G., and Teo, M. (2009). Institutional Investors, Past Performance, and Dynamic Loss Aversion. *Journal of Financial and Quantitative Analysis*, *44*(01), 155-188.

Pareto, C. Investopedia: Mutual Fund or ETF: Which is Right For You? Retrieved 3/1/15 from http://www.investopedia.com/articles/exchangetradedfunds/08/etf-mutual-fund-difference.asp.

Sirri, E. R. and Tufano, P. (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance*, *53*(5), 1589-1622.

Silvapulle, P., Moosa, I. A., and Silvapulle, M. J. (2004). Asymmetry in Okun's Law. *Canadian Journal of Economics*, 37(2), 353–374.

Tversky, A and Kahneman, D. (1974) Judgment Under Uncertainty: Heuristics and Biases. *Science*, *185* (4157), 1124-1131.

Woodward, G., and Anderson, H. M. (2009). Does Beta React to Market Conditions? Estimates of 'Bull' and 'Bear' Betas using a Nonlinear Market Model with an Endogenous Threshold Parameter. *Quantitative Finance*, 9(8), 913-924.

Zacks Funds (2014). 3 Thriving ETFs With Over 500% AUM Growth In 2014. Retrieved 2/28/15 from http://www.nasdaq.com/article/3-thriving-etfs-with-over-500-aum-growth-in-2014-etf-news-and-commentary-cm426825.

TABLE 1. LIST OF SECTOR ETF MUTUAL FUNDS

Sector Fund Name	Ticker	Inception Date	AUM[1]
Benchmark: Standard and Poor's Select Sector Indexes			
The Consumer Discretionary Select Sector SPDR Fund	XLY	12/16/1998	$8.95B
The Consumer Staples Select Sector SPDR Fund	XLP	12/16/1998	10.29B
The Energy Select Sector SPDR Fund	XLE	12/16/1998	12.68B
The Financial Select Sector SPDR Fund	XLF	12/16/1998	18.62B
The Health Care Select Sector SPDR Fund	XLV	12/15/1998	13.27B
The Industrial Select Sector SPDR Fund	XLI	12/15/1998	8.28B
The Materials Select Sector SPDR Fund	XLB	12/15/1998	3.07B
The Technology Select Sector SPDR Fund (includes telecommunications)	XLK	12/15/1998	13.26B
The Utilities Select Sector SPDR Fund	XLU	12/15/1998	7.29B
Benchmark : Dow Jones U.S. Sector Indexes			
iShares Dow Jones U.S. Consumer Services Sector Index Fund	IYC	6/11/ 2000	931.50M
iShares Dow Jones U.S. Consumer Goods Sector Index Fund	IYK	6/11/ 2000	663.2M
iShares Dow Jones U.S. Energy Sector Index Fund	IYE	6/12/ 2000	1.32B
iShares Dow Jones U.S. Financial Sector Index Fund	IYF	5/21/2000	1.33B
iShares Dow Jones U.S. Healthcare Sector Index Fund	IYH	6/12/ 2000	2.18B
iShares Dow Jones U.S. Industrial Sector Index Fund	IYJ	6/11/ 2000	869.78M
iShares Dow Jones U.S. Basic Materials Sector Index Fund	IYM	6/11/ 2000	520.96M
iShares Dow Jones U.S. Technology Sector Index Fund	IYW	5/14/2000	3.07B
iShares Dow Jones U.S. Telecommunications Sector Index Fund	IYZ	5/21/2000	514.66M
iShares Dow Jones U.S. Utilities Sector Index Fund	IDU	6/11/ 2000	1.52B
Bechmark: MSCI US Investable Markets Indexes			
Vanguard Consumer Discretionary ETF	VCR	1/25/2004	1.74B
Vanguard Consumer Staples ETF	VDC	1/26/2004	2.97B
Vanguard Energy ETF	VDE	9/22/2004	4.46B
Vanguard Financials ETF	VFH	1/26/2004	2.71B
Vanguard Health Care ETF	VHT	1/25/2004	5.55B
Vanguard Industrials ETF	VIS	9/22/2004	2.12B
Vanguard Information Technology ETF	VGT	1/25/2004	7.66B

Vanguard Materials ETF	VAW	1/26/2004	1.52B
Vanguard Telecommunications Services ETF	VOX	9/23/2004	989.26M
Vanguard Utilities ETF	VPU	1/25/2004	2.39B

TABLE 2. SUMMARY STATISTICS

This table presents summary statistics of the ETFs included in the sample. All figures are time series averages monthly data for each ETF. *Net flow* is net flow of fund calculated using the formula in the text and given in millions of dollars, *% Return* is the percent return on each ETF, *%Prem* is premium in percentage calculated as Premium = (Price – NAV)/NAV, *TNA* is total net assets in million dollars, *% Exp* is expense ratio in percentage of a fund's average net assets, *%TO* is total turnover ratio in percentage, *%Spread* is the bid–ask spread is calculated as spread=(ask-bid)/price and expressed in percentage.

	Net Flow ($m)		*% Return*		*% Prem*	*TNA (m$)*	*% Exp*	*% TO*	*% Spread*
SPDR									
Consumer Disc.	39.98	282.37	0.85	5.40	–0.003	1848.81	0.22	8.67	0.07
Consumer Staples	27.61	291.14	0.83	3.04	–0.007	3063.68	0.22	9.17	0.07
Energy	17.32	727.14	1.18	6.63	–0.012	5821.76	0.22	9.86	0.04
Financial	162.00	769.35	0.14	7.25	–0.006	6017.98	0.22	11.75	0.08
Health Care	33.89	264.85	0.79	3.88	–0.015	3143.98	0.22	5.31	0.06
Industrial	57.37	296.36	0.81	5.61	0.008	2462.17	0.22	6.03	0.07
Materials	20.85	194.98	0.81	6.25	–0.001	1684.96	0.22	12.44	0.08
Technology	67.21	303.51	0.72	5.05	–0.036	4811.98	0.22	7.25	0.08
Utilities	2.22	284.43	0.67	3.85	–0.054	3609.69	0.22	5.92	0.06
iShares									
Consumer Services	0.31	21.75	0.83	4.70	–0.029	249.30	0.51	6.69	0.09
Consumer Goods	–1.69	23.92	0.79	3.62	0.011	365.87	0.51	6.72	0.11
Energy	3.15	57.30	1.11	6.38	–0.018	922.82	0.51	7.47	0.12
Financial	9.81	100.30	0.23	6.66	0.032	530.26	0.51	9.47	0.09
Healthcare	2.29	68.84	0.83	3.88	0.001	930.45	0.51	5.50	0.10
Industrial	8.63	39.60	0.83	5.65	–0.026	425.94	0.51	5.33	0.11
Basic Materials	0.39	43.17	0.87	7.22	–0.026	609.85	0.50	8.86	0.10
Technology	15.30	63.42	0.75	5.55	0.707	1144.91	0.51	6.11	0.08
Telecommunications	–1.06	43.42	0.56	5.28	0.32	608.83	0.51	25.06	0.14
Utilities	–3.51	71.35	0.66	3.84	–0.034	682.89	0.51	6.50	0.11
Vanguard									
Consumer Disc.	8.43	34.51	0.88	5.68	–0.006	284.98	0.41	8.50	0.11
Consumer Staples	8.80	94.37	0.88	3.09	0.009	598.59	0.22	11.00	0.11

Energy	14.83	94.64	1.16	6.67	–0.005	1083.89	0.22	15.50	0.13
Financials	12.84	65.31	0.24	6.66	0.024	524.46	0.22	9.78	0.14
Health Care	13.98	92.87	0.85	3.92	0.017	662.50	0.22	8.72	0.10
Industrials	10.21	73.59	0.85	5.79	–1.841	341.69	0.22	8.61	0.12
Info. Technology	28.05	69.92	0.80	5.53	–0.009	1163.07	0.22	8.28	0.10
Materials	5.16	40.09	0.90	6.51	0.006	401.74	0.22	10.17	0.12
Telecommunications	3.10	39.95	0.75	4.78	–0.017	248.40	0.22	27.22	0.14
Utilities	8.09	98.38	0.70	3.82	0.005	553.95	0.22	9.11	0.10

TABLE 3. PANEL VAR RESULTS

This table presents the results of the unrestricted VAR in the level $Z_{it} = A_0+A_1 Z_{it-1}+f_i+U_{it}$, where, $Z_{it}\equiv[Flow_{it}, Ret_{it}, Prem_{it}]'$, A_0 and A_1 are vector of constant and slope coefficients, respectively, and f is individual ETF–specific fixed effects, and U_{it} is the vector of error terms. Details of the endogenous variables are given in the text. The sample consists of 29 sector ETFs divided into a full sample and 3 subsamples. The monthly data is from January 2005 to December 2013.

	Impact on					
Impact of						
	Panel A: Full Sample			**Panel B: SPDR**		
$Flow_{it-1}$	0.153* (1.82)	0.010** (2.84)	–0.001 (–0.53)	0.058 (0.64)	0.001 (0.31)	0.001 (0.94)
Ret_{it-1}	–0.175 (–0.78)	0.124** (4.10)	–0.002 (–1.21)	0.621 (1.29)	0.149** (2.90)	0.002 (–0.19)
$Prem_{it-1}$	0.106 (0.51)	0.968** (88.82)	0.175** (2.25)	6.255 (1.02)	–1.893* (–1.67)	0.046 (0.75)
$Flow_{it-2}$	–0.016 (–1.14)	–0.005** (–2.49)	–0.001 (–0.73)	0.011 (1.20)	–0.002** (–2.19)	0.001 (–0.17)
Ret_{it-2}	0.688** (3.90)	–0.033 (–1.51)	0.019** (2.15)	0.329 (1.05)	–0.066 (–1.48)	0.003 (1.59)
$Prem_{it-2}$	0.045 (0.12)	–0.101** (–3.22)	–0.082 (–1.12)	–4.295 (–1.26)	1.559 (1.52)	–0.063 (–1.15)
$Flow_{it-3}$	–0.017** (–2.34)	–0.010** (–2.92)	–0.001* (–1.86)	–0.005 (–0.64)	–0.006** (–2.04)	0.003 (0.08)
Ret_{it-3}	0.694** (2.50)	0.066** (2.83)	0.007 (0.82)	–0.246 (–0.61)	0.090* (1.91)	0.001 (–0.19)
$Prem_{it-3}$	–0.145 (–1.11)	–0.009 (–1.11)	0.028 (0.99)	2.765 (0.68)	–0.336 (–0.35)	–0.049 (–0.93)
	Panel C: iShares			**Panel D: Vanguard**		
$Flow_{it-1}$	0.173 (0.79)	–0.003 (–0.84)	–0.001 (–0.57)	0.206** (2.32)	0.021** (6.06)	–0.000 (–0.70)
Ret_{it-1}	0.174** (2.79)	0.114* (1.89)	–0.001 (–0.21)	–0.008 (–0.02)	0.113** (2.45)	–0.002 (–1.14)

$Prem_{it-1}$	0.143 (0.83)	0.996** (143.27)	0.182** (2.35)	3.613 (0.23)	0.658 (0.53)	0.018 (0.36)
$Flow_{i2}$	0.009 (0.30)	–0.001 (–0.15)	–0.011** (–4.29)	–0.070** (–2.35)	–0.009** (–3.12)	0.001** (2.99)
Ret_{it-2}	–0.002 (–0.02)	–0.013 (–0.29)	0.003 (0.94)	1.616** (4.31)	–0.025 (–0.69)	0.002* (1.66)
$Prem_{it-2}$	–0.355 (–1.05)	–0.109* (–1.79)	–0.101 (–1.43)	14.050 (1.25)	0.514 (0.54)	–0.010 (–0.22)
$Flow_{it-3}$	–0.063** (–2.89)	–0.001 (–0.53)	–0.004 (–1.14)	–0.054** (–2.54)	–0.014** (–3.26)	–0.000 (–1.48)
Ret_{it-3}	–0.182* (–1.77)	0.037 (1.18)	0.021 (0.85)	2.384** (3.89)	0.066* (1.67)	–0.001 (–1.26)
$Prem_{it-3}$	0.240 (0.57)	0.007 (0.14)	0.156** (1.98)	–0.050** (–2.25)	–0.016** (–7.79)	–0.001** (–20.15)

Note: Three variable panel VAR as in equation (1) is estimated by GMM, where fixed effects are removed prior to estimation. The optimum lag of the VAR system is chosen to minimize Modified AIC. (MAIC). Reported numbers show the coefficients of regressing the column variables on the row variables. Heteroskedasticity adjusted *t*–statistics are reported in parentheses. * (**) denotes significance at 10% (5%) level.

TABLE 4. RETURN, FLOW AND SENTIMENT: FORECAST-ERROR VARIANCE DECOMPOSITION

	Impact on					
Impact of	**Full Sample**			**SPDR**		
$Flow_{it}$	98.25%	1.85%	0.05%	99.47%	0.51%	0.04%
Ret_{it}	1.57%	86.34%	0.34%	0.45%	98.54%	0.35%
$Prem_{it}$	0.18%	11.81%	99.61%	0.09%	0.95%	99.61%
	iShares			**Vanguard**		
$Flow_{it}$	99.52%	0.43%	0.77%	88.63%	6.69%	2.99%
Ret_{it}	0.29%	64.21%	0.10%	11.15%	93.23%	3.38%
$Prem_{it}$	0.19%	35.37%	99.13%	0.22%	0.08%	93.62%

Note: Variation in a row variable explained by a column variable. The numbers in percentage represent accumulated variances over 10 periods.

TABLE 5. PANEL VAR-GRANGER CAUSALITY WALD TEST

Equation	Excluded	Full Sample	SPDR	iShares	Vanguard
$Flow_{it}$	Ret_{it}	19.79 (0.000)**	5.33 (0.149)	15.67 (0.001)**	24.74 (0.000)**
	$Prem_{it}$	2.03 (0.566)	1.73 (0.631)	1.88 (0.597)	6.60 (0.086)*
	All	20.14 (0.003)**	6.13 (0.409)	18.10 (0.006)**	28.49 (0.000)**
Ret_{it}	$Flow_{it}$	25.93 (0.000)**	11.23 (0.011)**	1.11 (0.776)	66.47 (0.000)**
	$Prem_{it}$	36.76 (0.000)**	4.34 (0.227)	54.40 (0.000)**	87.11 (0.000)**
	All	89.64 (0.000)**	14.47 (0.025)**	27.72 (0.000)**	56.44 (0.000)**
$Prem_{it}$	$Flow_{it}$	4.23 (0.238)	0.99 (0.804)	26.35 (0.000)**	10.15 (0.017)**
	Ret_{it}	5.37 (0.147)	2.58 (0.461)	4.29 (0.231)	6.14 (0.105)
	All	6.62 (0.357)	3.34 (0.765)	31.97 (0.000)**	14.71 (0.023)**

Note: denotes the Wald test statistics under the null hypothesis that the excluded variable(s) does not Granger-cause equation variable. Associated is reported in the parentheses. * (**) denotes significance at 10% (5%) level.

TABLE 6. PANEL VAR WITH ASYMMETRIC RESPONSES

This table presents the results of the unrestricted VAR in the level, , as described in Table 3. The only difference is in the vector of endogenous variables which now includes three more variables, , where additional variables are used to capture asymmetric responses. See Appendix for further details.

	Impact on					
Impact of						
	Panel A: Full Sample			**Panel B: SPDR**		
$Flow_{it-1}$	0.037	–0.047**	–0.001	–0.138**	–0.049*	–0.001
	(0.59)	(–3.53)	(–0.32)	(–2.00)	(–1.86)	(–0.36)
$Flow^{-}_{it-1}$	0.112	0.059**	0.001	0.204*	0.052**	0.001
	(0.93)	(4.11)	(0.29)	(1.82)	(1.98)	(0.41)
Ret_{it-1}	0.419	–0.034	0.014	–0.998	0.085	–0.003
	(0.72)	(–0.65)	(1.37)	(–0.75)	(0.94)	(–0.85)
Ret^{-}_{it-1}	–0.886	0.342**	–0.033	0.991	0.067	0.003
	(–1.11)	(3.16)	(–1.50)	(0.61)	(0.38)	(0.41)
$Prem_{it-1}$	1.704**	0.949**	–0.087	7.792	–6.129**	–0.146
	(3.68)	(22.74)	(–0.48)	(1.12)	(–3.12)	(–1.10)
$Prem^{-}_{it-1}$	–2.835**	0.009	0.476	–1.724	6.074**	0.293*
	(–3.36)	(0.14)	(1.52)	(–0.10)	(2.28)	(1.72)
$Flow_{it-2}$	0.083	–0.020*	–0.004	0.118	–0.013	0.000
	(1.10)	(–1.73)	(–1.07)	(1.48)	(–0.66)	(–0.48)
$Flow^{-}_{it-2}$	–0.105	0.015	0.004	–0.114	0.011	0.000
	(–1.40)	(1.24)	(0.95)	(–1.39)	(0.56)	(0.50)
Ret_{it-2}	0.528**	0.035	–0.001	0.306	0.091	0.001
	(2.12)	(0.75)	(–0.13)	(1.17)	(1.21)	(0.34)
Ret^{-}_{it-2}	0.434	–0.172**	0.034	0.020	–0.396**	0.003
	(0.83)	(–2.16)	(1.03)	(0.05)	(–2.65)	(0.47)
$Prem_{it-2}$	0.201	0.057	–0.149	4.084	–2.129	–0.049
	(0.17)	(0.92)	(–1.01)	(0.85)	(–1.14)	(–0.38)
$Prem^{-}_{it-2}$	–0.084	–0.313**	0.083	–12.148	5.176**	–0.037
	(–0.05)	(–2.62)	(0.28)	(–1.46)	(2.02)	(–0.24)
	Panel C: iShares			**Panel D: Vanguard**		
$Flow_{it-1}$	0.025	–0.048**	–0.016	0.095	–0.019	0.001
	(0.30)	(–2.04)	(–0.66)	(0.61)	(–1.44)	(1.11)
$Flow^{-}_{it-1}$	0.163	0.050**	0.016	0.072	0.041**	–0.001
	(0.49)	(2.06)	(0.65)	(0.43)	(3.01)	(–1.32)
Ret_{it-1}	0.019	–0.094	–0.001	2.267**	–0.088	–0.004*
	(0.12)	(–0.87)	(–0.08)	(2.87)	(–1.18)	(–1.65)
Ret^{-}_{it-1}	0.332	0.424*	0.002	–3.755**	0.402**	0.003
	(1.33)	(1.78)	(0.07)	(–3.36)	(2.56)	(0.70)
$Prem_{it-1}$	0.898**	1.028**	–0.028	–9.445	–4.955**	–0.029
	(2.51)	(39.20)	(–0.15)	(–0.36)	(–2.33)	(–0.24)

$Prem^{-}_{it-1}$	–1.357**	–0.073	0.388	15.984	9.943**	0.087
	(–2.57)	(–1.48)	(1.26)	(0.46)	(3.05)	(0.58)
$Flow_{it-2}$	0.046	–0.016	–0.032	–0.093	–0.017	0.000
	(0.57)	(–0.74)	(–1.13)	(–0.69)	(–0.93)	(0.70)
$Flow^{-}_{it-2}$	–0.046	0.015	0.024	0.036	0.006	0.000
	(–0.52)	(0.68)	(0.81)	(0.27)	(0.34)	(0.37)
Ret_{it-2}	–0.276	–0.004	0.013	1.077*	0.061	–0.002
	(–1.40)	(–0.04)	(0.59)	(1.73)	(0.89)	(–1.07)
Ret^{-}_{it-2}	0.480	–0.083	–0.018	1.411	–0.218*	0.007
	(1.52)	(–0.46)	(–0.40)	(1.19)	(–1.75)	(1.52)
$Prem_{it-2}$	–0.347	0.142	–0.064	54.014**	–0.442	–0.033
	(–0.41)	(1.25)	(–0.47)	(2.10)	(–0.21)	(–0.47)
$Prem^{-}_{it-2}$	0.079	–0.467*	–0.096	–67.700**	1.541	0.044
	(0.07)	(–1.94)	(–0.41)	(–2.08)	(0.56)	(0.38)

Note: See Table 3

Exhibit 1

This appendix outlines the asymmetric response model employed in this paper to report results in Table 6. The VAR specification in the level, $Z_{it} = A_0+A_1 Z_{it-1}+f_i+U_{it}$, is used as in Table 3. The only difference is in the vector of endogenous variables which now includes three more variables, $Z_{it} = [Flow_{it}, Flow^{-}_{it}, Ret_{it}, Ret^{-}_{it}, Prem_{it}, Prem^{-}_{it}]'$, where, $Flow^{-}_{it} = Min(Flow_t,0)$, $Ret^{-}_{it} = Min(Ret_t,0)$ and $Prem^{-}_{it} = Min(Prem_t,0)$. Using an example of the relationship between and , we show the basis of the asymmetric model is used in this paper.

Consider a simple regression specification that can capture asymmetric relationship between the dependent variable, $Flow_t$, and right hand side variable return, Ret_t,

$$Flowt = \alpha+\beta^{+}Ret_t^{+}+\beta^{-}Ret_t^{-}+e_t \quad \text{(A1)}$$

where, $Ret_t^{+} = Max(Ret_t,0)$ and $Ret_t^{-} = Min(Ret_t,0)$ are positive and negative values on the independent variable, respectively. Similarly, β^{+} and β^{-} are coefficients associated with positive and negative values of Ret_t, respectively. Recall that $Ret_t = Ret_t^{+}Ret_t^{-}$, so that $Ret_t^{+} = Ret_t - Ret_t^{-}$ which can be replaced in A1 to get,

$$Flowt = \alpha+\beta^{+}(Ret_t+Ret_t^{-})+\beta^{-}Ret_t^{-}+e_t \quad \text{(A2)}$$

Opening the parentheses,

$$Y_t = \alpha+\beta^{+}Ret_t - \beta^{+}Ret_t^{-} + \beta^{-}Ret_t^{-} +e_t \quad \text{(A3)}$$

and, rearranging we get,

$$Y_t = \alpha+\beta Ret_t + \beta^{+}Ret_t^{-} + \delta\, Ret_t^{-} +e_t \quad \text{(A4)}$$

where, β replaces β^{+} for simplicity and $\delta = (\beta^{-} - \beta^{+})$. Therefore, the impact of separate positive and negative values will be judged by the significance of the coefficient, δ. Under the null hypothesis, $\delta = 0$ implies $\beta^{-} = \beta^{+}$, indicating no difference between the positive and negative beta. Similar derivation applies to other variables in the asymmetric specification.

Note: For similar formulation, see Silvapulle et al. (2004, p. 362), and Woodward and Anderson (2009, p.916).