ETFs and Their Impact on Bid-Ask Spread

Achyut Gautam
Ramapo College of New Jersey, Mahwah, NJ, 07430

INTRODUCTION

The growth in passive fund investment has been steadily increasing over the last two decades. Within the US, passive funds grew by 138% from 2007-2017 and have about 15% share of total outstanding securities (Sushko et al.). Exchange Traded Fund, commonly known as ETF, is a type of passive fund that tracks a stock index, commodities, bonds, or a basket of assets.
At the end of September 2017, global ETF assets totaled to $4.4 trillion with a growth rate 21% from 2015 (Julie et al.). The U.S. is the world's largest ETF market and as of 2017, the U.S. ETF industry alone had roughly $3 trillion in total industry asset.

Previous literature have found that an increase in ETF trades of a particular stock leads to an increase in trading cost which is captured by an increase in bid-ask spread of the underlying (Israeli et al.). Similarly, stocks that are traded as ETFs reflect higher volatility than otherwise similar securities (Ben-David et al.).

Since ETFs are easier to trade and are cost-effective, ETFs encourage passive investment leading to lesser scrutiny from the market. Based on this rationale, the research expands on previous literatures by finding the incremental change in bid-ask spread, a transaction cost, of entire stock market index when volume of ETF tracking the index increases.

THEORETICAL BACKGROUND

$$HLSPREAD = \beta_0 + \beta_1 \%\Delta ETF_m + \varepsilon \quad (1)$$
$$HLSPREAD = \beta_0 + \beta_1 \%\Delta ETF_m + \beta_2 \%\Delta ETF_n + \beta_3 Volatility + \varepsilon \quad (2)$$
$$HLSPREAD_{CS} = \beta_0 + \beta_1 \%\Delta ETF_m + \beta_2 \%\Delta ETF_n + \beta_3 Volatility + \varepsilon \quad (3)$$

HLSPREAD - Monthly high-low measure of bid-ask spread for an index over a month
$HLSPREAD_{CS}$ - Corwin and Schultz (2012) annual high-low a measure of bid-ask spread.
Δ % ETF - Monthly change in the percentage of ETF traded
Volatility - The CBOE Volatility Index (VIX) for and the CBOE Nasdaq Volatility Index (VXN) are used as proxies for the volatility.

High-Low Spread Measure

High – Low spread in in Model 1 and 2 is obtained in the following way:

$$HLSPREAD_{i,t} = \frac{HIGH(i,t) - LOW(i,t)}{LOW(i,t)}$$

where *i* represents the stock market index and *t* represents a time period.

High – low spread in Model 3 is obtained in the following way:

$$S = \frac{2(e^{\alpha}-1)}{1+e^{\alpha}}, \quad \alpha = \frac{\sqrt{2\beta}-\beta}{3-2\sqrt{2}} - \sqrt{\frac{\delta}{3-2\sqrt{2}}}, \quad \beta = \sum_{j=0}^{i}\left[\ln\left(\frac{H^{\circ}_{t+j}}{L^{\circ}_{t+j}}\right)\right]^2, \quad \delta = \left[\ln\left(\frac{H^{\circ}_{t,t+1}}{L^{\circ}_{t,t+1}}\right)\right]^2$$

where S = Corwin-Schultz bid-ask spread estimator.

Rationality predicts that increase in ETFs trades increase bid-ask spread of the underlying. Volatility, is expected to have positive relationship as during a rapid market change bid-ask spread is much wider as market makers have greater opportunities to take advantage of it.

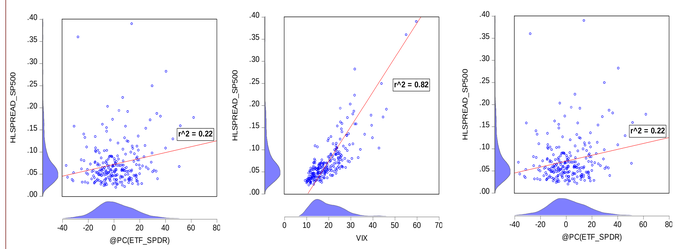
CLASSICAL LINEAR ASSUMPTIONS

The model satisfies five of the seven classical linear assumptions:

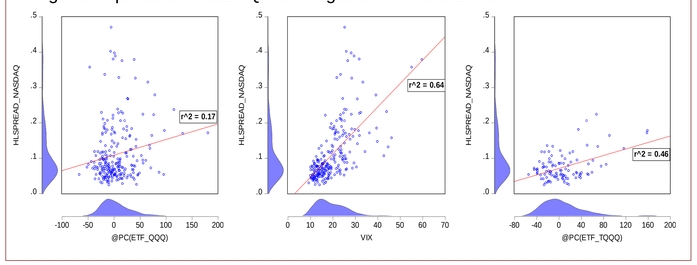
- ✓ The regression model is linear in the coefficients and the error term
- ✓ The error term has a population mean of zero
- ✓ All independent variables are uncorrelated with the error term
- ✓ Observations of the error term are uncorrelated with each other
- ✓ No independent variable is a perfect linear function of other explanatory variables
- χ The error term has a constant variance (no heteroscedasticity)
- χ The error term is normally distributed

SCATTER PLOT

High-low spread of S&P 500 against VIX and ETFs



High-low spread of NASDAQ – 100 against VXN and ETFs



REFERENCES

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. "Do ETFs increase volatility?." *The Journal of Finance* 73.6 (2018)
Corwin, Shane A., and Paul Schultz. "A simple way to estimate bid-ask spreads from daily high and low prices." *The Journal of Finance* 67.2 *(2012)*
Grant, Turner, and Vladyslav Sushko. "The implication of passive investing for securities markets." *The Journal of Finance* 73.6 (2018)
Sushko, Vladyslav, and Grant Turner. "The Implications of Passive Investing for Securities Markets." *The Bank for International Settlements*, 11 Mar. 2018.
Glosten, Lawrence, Suresh Nallareddy, and Yuan Zou. *ETF trading and informational efficiency of underlying securities*. Working paper. Columbia University. October, 2015.
Gujarati, Damodar N., and Dawn C. Porter. *Basic Econometrics*. McGraw-Hill/Irwin, 2017
Israeli, Doron, Charles MC Lee, and Suhas A. Sridharan. "Is there a dark side to exchange traded funds? An information perspective." *Review of Accounting Studies* 22.3 (2017)
Julie, Kerr, Lisa Kealy, and Matt Forstenhausler. *Global ETF Research*. Ernst & Young Global Limited.
Murphy, Cinthia. "What US ETF Market Looks Like Today." *ETF.com*

REGRESSION RESULTS

Sample (adjusted): 2000M06 2018M12
Included observations: 223 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.030777	0.004799	-6.412817	0.0000
@PC(ETF_SPDR)	0.000250	0.000126	1.977495	0.0492
@PC(ETF_IVV)	0.000146	4.11E-05	3.544833	0.0005
VIX	0.005155	0.000229	22.49610	0.0000
R-squared	0.730813	Mean dependent var		0.072232
Adjusted R-squared	0.727126	S.D. dependent var		0.051891
S.E. of regression	0.027106	Akaike info criterion		-4.360323
Sum squared resid	0.160911	Schwarz criterion		-4.299208
Log likelihood	490.1760	Hannan-Quinn criter.		-4.335651
F-statistic	198.1874	Durbin-Watson stat		2.144074
Prob(F-statistic)	0.000000			

Table 1: hslpread c spdr ivv vix

Sample (adjusted): 2000M10 2018M12
Included observations: 219 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.021188	0.005099	-4.155184	0.0000
@PC(ETF_QQQ)	0.000237	7.75E-05	3.062925	0.0025
VXN	0.004739	0.000172	27.58364	0.0000
R-squared	0.786956	Mean dependent var		0.103955
Adjusted R-squared	0.784984	S.D. dependent var		0.078207
S.E. of regression	0.036264	Akaike info criterion		-3.782359
Sum squared resid	0.284062	Schwarz criterion		-3.735934
Log likelihood	417.1683	Hannan-Quinn criter.		-3.763609
F-statistic	398.9383	Durbin-Watson stat		1.823475
Prob(F-statistic)	0.000000			

Table 2: hslpread c qqq tqqq vxn

Sample (adjusted): 2010M03 2018M12
Included observations: 106 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.020434	0.006560	-3.114803	0.0024
@PC(ETF_QQQ)	7.66E-05	8.12E-05	0.943911	0.3474
@PC(ETF_TQQQ)	-0.000218	7.87E-05	-2.765313	0.0068
VXN	0.002061	0.000337	6.115402	0.0000
R-squared	0.277078	Mean dependent var		0.017857
Adjusted R-squared	0.255816	S.D. dependent var		0.020154
S.E. of regression	0.017386	Akaike info criterion		-5.229251
Sum squared resid	0.030833	Schwarz criterion		-5.128744
Log likelihood	281.1503	Hannan-Quinn criter.		-5.188515
F-statistic	13.03137	Durbin-Watson stat		1.751386
Prob(F-statistic)	0.000000			

Table 3: hslpread$_{cs}$ c spdr ivv vix

Sample (adjusted): 2010M03 2018M12
Included observations: 106 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.020434	0.006560	-3.114803	0.0024
@PC(ETF_QQQ)	7.66E-05	8.12E-05	0.943911	0.3474
@PC(ETF_TQQQ)	-0.000218	7.87E-05	-2.765313	0.0068
VXN	0.002061	0.000337	6.115402	0.0000
R-squared	0.277078	Mean dependent var		0.017857
Adjusted R-squared	0.255816	S.D. dependent var		0.020154
S.E. of regression	0.017386	Akaike info criterion		-5.229251
Sum squared resid	0.030833	Schwarz criterion		-5.128744
Log likelihood	281.1503	Hannan-Quinn criter.		-5.188515
F-statistic	13.03137	Durbin-Watson stat		1.751386
Prob(F-statistic)	0.000000			

Table 4: hslpread$_{cs}$ c qqq tqqq vxn

CONCLUSION

The study concludes that there is a strong significant positive relationship between the bid-ask spread of S&P 500 and its ETFs (Table 1) while the same holds true but only for NASDAQ-100 and QQQ (Table 2). TQQQ had relatively small observation, failed the omitted variable test, and led to insignificant result. Thus it was left out from the equation 2. Model 3, which uses Corwin and Schultz spread, resulted in either wrong coefficient sign or insignificant result (Table 3, 4).

Unlike the model used by Israeli et al., this study paper excludes control variables such as institutional ownership and number of analysts analyzing the underlying due to lack of data. Incorporating control variables and effectively controlling for noise might have increased the robustness and efficiency of the models, established homoscedasticity and normality assumption, and improved results of the regression model.