Time Varying Price Discovery in VIX Exchange Traded Notes: A Tale of Retail vs. Institutional Trades.

Bart Frijns Ilnara Gafiatullina Alireza Tourani-Rad
Adrian Fernandez-Perez

Auckland University of Technology
Auckland, New Zealand

August 4, 2016

Abstract

This study investigates the intraday price discovery between the VIX short-term futures ETNs (VXX) and inverse VIX short-term ETNs (XIV) for the period January 3, 2011 to December 31, 2015. Applying both the Hasbrouck (1995) information share and the Gonzalo and Granger (1995) common factor weight approach, we observe strong time variation in the price discovery contribution between the direct and inverse notes. We find that the classical measures of trading costs and market liquidity are significant determinants of price discovery. We also document that price discovery of the VXX increases with the greater institutional ownership and on those days the level of the VIX tends to be high. We also report that order imbalance of both retail and institutional investors has a significant effect on price discovery.

Keywords: Volatility ETNs, High-Frequency Data, Price Discovery, Institutional Ownership, Order Imbalance.

1 Introduction

The VIX futures exchange traded notes (ETNs) were first issued in 2009 by Barclays Bank. During the next year Credit Suisse and UBS designed and issued new series of volatility ETNs. These new products have gained a huge popularity among market participants, currently there are 26 VIX related ETPs.[1] These products allow market participants to take a direct, indirect or leveraged positions in the S&P 500 volatility with static or dynamic allocation between VIX short-term and medium-term VIX futures contracts. The most popular direct ETN is the iPath S&P 500 VIX Short-Term Futures ETN (the VXX) and inverse ETN is the VelocityShares Daily Inverse VIX Short Term ETN (the XIV), with an average trading daily volume being 70 and 30 mln contracts by the end of 2015, respectively. These series of ETNs are short-term notes that are uncollaterized debt securities which are benchmarked to the similar indices which reflect the returns potentially available through investing in a long position of the VIX futures contract with a constant maturity of one month. The major difference between the two ETNs is that inverse ETN promises the daily return which is the inverse to the daily return of an underlying index. Despite such a popularity, the VXX lost about 99 % of its value and underwent three reverse splits since the inception. So what makes these contracts so attractive?

Investors have always been trying to limit their exposure to the downside risk without sacrificing the potential for the upside gains. The observed changes in the VIX are negatively correlated with the changes in the S&P 500 index, thus adding volatility exposure to the portfolio can be potentially a risk mitigation strategy during market downturns (Signori (2010)). This property drove interest in volatility related products as an asset class which could provide an effective hedge during market shortfalls. However, several studies conclude that the VIX futures ETNs do not provide effective hedges when held as a passive buy-and-hold

[1] An exchange-traded product (ETP) is a derivative security which is traded on an exchange. ETPs are typically benchmarked to indices, stocks, commodities, or may be actively managed. There are several different types of ETPs, including Exchange-traded funds (ETFs) and Exchange-traded notes (ETNs).

investment, primarily because the long-term performance of the VIX and the VIX futures ETNs shows substantial deviation which is due to the negative roll-over costs during contango markets (Alexander and Korovilas (2012*a*), Whaley (2009)).

While a lot of studies have focused on considering volatility as an asset class and on its diversification benefits, a few addressed the question of informational dominance between the VIX and its derivatives, or between the different volatility products. Shu and Shang (2012), for instance, examine the lead-lag dynamics between the VIX and VIX futures prices on the daily basis and conclude that VIX futures market leads the spot VIX. To avoid issues related to the data aggregation, Frijns et al. (2015) use intraday data and find the evidence of bi-directional causality between those two markets with the VIX futures prices have become more informative over the time. As for volatility ETPs, Bornardo et al. (2016) study the price discovery relationship within direct, leveraged and inverse VIX ETPs employing 1-min data and identify the price discovery leader within each category. However, none of the research so far has addressed the price discovery function between different ETNs categories which might be attractive for informed traders during different market volatility regimes and thus exhibit switching pattern in informational efficiency.

Traditional market microstructure literature distinguishes two categories of traders according to their possession of fundamental information about assets' prices. They are informed traders and uninformed or noise traders. New information becomes impounded in prices and revealed to the market through the trading activities of informed traders. If informed traders are more likely to choose one particular market to transact then this market dominates the price discovery function and tends to lead prices of other market. Whaley et al. (2014) point out that direct and inverse ETPs have different investor structures. On the reporting day the VXX was largely held by retail customers and the XIV was dominated by institutions. Given the tremendous trading volume in the VXX does that mean that it is made up only from potentially noise trading which is associated with lower informational efficiency of prices? Considering institutions as a better informed traders relative to the retailers can we conclude that informational content of the XIV prices is always higher than those of the VXX? Or maybe there is some time variation in trading pattern among the institutions due to the changing market conditions and subsequently some time variation in informational efficiency between the VXX and the XIV?

In this paper we investigate the price discovery between the VXX and XIV ETNs using hight frequency data. We employ two popular approaches which are widely used to measure the price discovery contribution: the common factor model (CS) introduced by Gonzalo and Granger (1995) and the information shares (IS) model developed by Hasbrouck (1995). Our results show strong time variation in price discovery between the VXX and the XIV. We also study the key determinants of price discovery and relate our findings with the institutional ownership which is calculated as the institutional shares held, taken from the quarterly 13-F reports, divided by shares outstanding. The final step of our analysis considers the effect of order imbalance of small and large trades on price efficiency.

This study contributes to the existing literature by providing empirical examination which volatility ETN plays the dominant role in the mechanism of price discovery in the VIX. Our findings help to understand the behavior of institutions and their imput to the efficient pricing of the volatility ENTs.

The remainder of this paper is organized as follows: Section 2 provides background information on the VIX and the two most popular direct and inverse volatility ETNs. In Section 3 we describe methodology adopted for this study. Section 4 explains the data used in this paper and summary statistics. In Section 5 we present the empirical results. Section 6 concludes.

2 Background

In this section, we are going to review some related literature, discuss pricing methodology and properties of the VIX as well as of the VXX and XIV. We also show that by the design of the underlying indices of these ETNs there should exist a cointegrating relationship between the VXX and the XIV with a vector $(1, 1)$.

2.1 Literature Review

Taking long position in implied volatility is a well documented strong diversification tool (Dash and Moran (2005), Daigler and Rossi (2006)). With the advent of the VIX in 1993, investors potentially could achieve timely protection of their equity portfolios during market downturns. However, the VIX itself is not a tradable asset but just a calculated number. To allow trading in volatility the VIX futures were introduced in 2004 and the VIX options in 2006. Trading in VIX derivatives might be too sophisticated for retail investors and many institutions might be restricted from buying futures and options contracts (Whaley (2013)). Thus the third generation of volatility products - the VIX futures ETPs - were introduced in 2009.

A number of researchers consider the possibility of hedging the portfolios with VIX futures and VIX futures ETNs and conclude that for passive buy-and-hold investors, the direct VIX futures ETNs are not effective hedges.

VIX futures ETNs are structured in a way that most of them do not provide the performance of the VIX but instead track the performance of a constant maturity futures index. Several studies attribute the substantial deviation between the long-term performance of the VIX spot index and the direct short-term futures ETNs to the roll-over loss which is associated with the term structure of futures market (see Husson and McCann (2011), Deng et al. (2012), Alexander and Korovilas (2012a), (2012b), DeLisle et al. (2014)). The reference index of short-term futures notes measures the return to a portfolio of one and two month VIX futures contracts which are rebalanced daily to achieve an average maturity of one month, thus being an indirect investment in the VIX. During the majority of the days the futures prices exhibit an upward-sloping term structure and only during market instability change to a downward-sloping term structure. The daily rebalancing implies selling a fraction of holdings in the one month futures contracts at a lower price than the price which was paid when it was purchased as a two month contract. This incurs a loss of value of the underlying index and, as a result, the value of the direct short-term futures notes, and in particular the VXX, shows a steady loss.

Bahaji and Aberkane (2015) go beyond the buy-and-hold strategy and show that rational risk-averse agent can enhance the performance of their portfolio by dynamically taking short or long position in the VIX futures.

The conflicting evidence of poor long-term hedging performance, decreasing prices of the VXX and enormous trading volume in this product naturally prompts us to question the informational content of the VXX prices. As Alexander and Korovilas (2012b) point out only during periods of market instability does the VIX futures term structure swing to backwardation and the VXX stops suffering from the negative roll cost effects thus being effective diversifiers/hedgers. The XIV is almost a mirror reflection of the VXX and it experiences a stable growth when the VXX falls and loses its value when the VXX goes up. That might materialize in a switching behavior of informed traders between the VXX and the XIV during different market regimes and hence to the flow of the efficiency between the above mentioned products.

2.2 VIX

The Chicago Board Option Exchange (CBOE) introduced the VIX in 1993. The VIX was designed to measure the market's expectation of 30-day volatility implied by the at-the-money S&P 100 Index option prices. The CBOE, together with Goldman Sachs, revised the VIX calculation in 2003. The new VIX is based on the S&P 500 Index (SPX), the core index for U.S. equities, and estimates expected volatility by averaging the weighted prices of SPX out-of-the money puts and calls over a wide range of strike prices.

The VIX formula is an adaptation of the results presented by Demeterfi et al. (1999) to the discrete market data. The VIX index value represents an annualized measure expressed in percentage points of 30-day expected volatility of the SPX. A synthetic option that expires in 30 days is calculated using the bid-ask prices of near and next-term options available with more than 23 days and less than 37 days to expiration. Thus the components of a synthetic 30-day option are "standard" and "weekly" SPX options. The generalized formula used in the VIX calculation described in the CBOE white paper [2] is:

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}[\frac{F}{K_0} - 1]^2 \tag{1}$$

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2} \tag{2}$$

$$F = e^{RT}(C(K,T) - P(K,T)) + K \tag{3}$$

where T is time to expiration measured as the fraction of numbers of minutes until the option maturity and the number of minutes in one year, F is forward index level desired from index option prices, K_0 is first strike below the forward index level, K_i is strike price of the ith out-of-the-money option; a call if $K_i > K_0$; and a put if $K_i < K_0$; both put and call if $K_i = K_0$, ΔK_i is interval between strike prices - half the difference between the strike on either side of K_i, R is risk-free interest rate to expiration and $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike K_i. The forward price in Equation (3) is derived via the put-call parity relation using a pair of put and call options with prices that are closest to each other. The detailed explanation of the current general formula 1 can be found in Carr and Wu (2006).

As the VIX calculation involves the use of near and next-term put and call options, in practice the calculation will consist of two risk-free interest rates R_1 and R_2,

[2]see the white paper of the VIX, available at http://www.cboe.com/micro/vix/vixwhite.pdf

two times to expiration T_1 and T_2, two strike prices at which the absolute difference between the call and put prices is smallest and, respectively, two different forward index prices F_1 and F_2. To calculate 30-day expected variance the CBOE interpolates between σ_1^2 and σ_2^2 calculated by Equation (1) for near and next-term available options, after that this value is annualized. The VIX is calculated as 100 times the square root of that value:

$$VIX = 100 \times \sqrt{\left\{T_1\sigma_1^2\left[\frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}}\right] + T_1\sigma_1^2\left[\frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}}\right]\right\} \times \frac{N_{365}}{N_{30}}}$$

The CBOE publishes near-term and next-term VIX "components" every 15 seconds during each CBOE trading day.

Historically, the VIX has an inverse relationship with the S&P 500 index. The correlation coefficient between the daily returns of the VIX and the SPX is about -0.7 for the period 1990-2015 and -0.85 for period 2011-2015. As a volatility measure the VIX should reflect the market movements in both direction, up or down, but when investors expect the downside market move they seek the portfolio insurance what results in increased demand for put options that in turn results in higher implied volatility. However, this is not the case during the market upward movements. Indeed, Bollen and Whaley (2004) find that implied volatility of S&P 500 options is mostly affected by buying pressure for index puts. Thus the VIX peaks during periods of market distress and that is why often referred to as "investor fear gauge". The asymmetry of the movements during the market's fall and rise is documented by Whaley (2009). Shu and Zhang (2012), DeLisle et al. (2014) also note that the VIX returns appears to be strongly negatively correlated with the S&P 500 price changes during large downside moves.

2.3 Volatility Exchange Traded Notes: VXX and XIV

In 2009, Barclays Bank issued the iPath S&P 500 VIX Short-Term Futures ETNs, or the VXX. [3] The return on the VXX is linked to the performance of the relevant

[3] see prospectus of the VXX at http://www.ipathetn.com/US/16/en/contentStore.app?id=5149530

index. The index is designed to reflect the returns potentially available through investing in a long position of the VIX futures contract with a constant weighted average maturity of one month, plus interest accrual on the notional value of the index based on the three-month U.S. Treasury rate. Thus the return of the index underlying the VXX contract consists of two components and bears the name Total Return (TR) index. The ETN performance is linked to the performance of the index underlying the ETN less an investor fee.

In 2010, Credit Suisse designed the VelocityShares Inverse VIX Short Term ETNs with exchange ticker XIV. [4] The inverse ETNs where created to allow investors to take a short position in the underlying index. The index underlying the XIV is almost identical to the one underlying the VXX with the exception that it excludes the U.S. three-month Treasury rate return component and bears the name Excess Return (ER) index .

The closing indicative value (CIV) of each of the ETNs is calculated by the issuers - Barclays and Credit Suisse on a daily basis. The value on the inception date was $100. The CIV of the ETN is designed to approximate the economic value of this ETN. At the prospectus the issuers warn that the observed trading prices of the ETNs might differ from the intraday indicative values and from the closing indicative values due to imbalances of supply and demand, lack of liquidity or the issuer's credit rating. The closing indicative values of both the VXX and the XIV are based on the closing level of the applicable underlying indices but have minor differences in their definitions. As a result, both of the closing indicative values include the U.S. three-month Treasury rate return component and have either direct or inverse relationship with the return of rolling futures position in the VIX index. The mathematical part of the explanation is presented in the next section.

Whaley (2013) and Bordonado et al. (2016) evaluate the tracking performance of the most popular VIX ETPs by comparing the daily returns of market prices

[4] http://app.velocitysharesetns.com/files/prospectus/ VelocityShares_ETN_Amended_Final_Pricing_Supplement_ VIX_AR40_long_form_2.PDF

with the daily returns of their respective benchmarks and by comparing the observed market prices with the closing indicative values. They conclude that both the VXX and the XIV closely, though not perfectly, track the underlying benchmarked indices and closing indicative values.

Whaley (2013) provides additional evidence that, although the short-term index returns have a negative correlation with stock returns, due to the contango trap direct VIX notes benchmarked to the short-term futures index are almost doomed to loose value and thus are not suitable buy-and-hold investment. However, as the mid-term indices suffer less from the contango trap because the futures curve gets less steep on the higher maturities and as the index's returns are positively correlate with the VIX index he suggests that they must be a better long term diversification tool. Also he suggest that while direct VIX ETNs depreciate in value the inverse ETNs appreciate in value and might be considered as a good buy-and-hold investment, though by theirs nature they do not bring diversification power.

2.3.1 Calculation of the Indices

The VXX is benchmarked to the S&P 500 VIX Short-Term Futures Index Total Return (TR), or SPVXSTR index. The value of the SPVXSTR index depends on its value on the previous day and two returns components. The first component measures the return from a rolling long position in the first and second month VIX futures contracts which creates a weighted average time to maturity of one month. The second component of the index's return includes interest accruals based on the 3-month U.S. Treasury Bill rate.

The XIV is benchmarked to the S&P 500 VIX Short-Term Futures Index Excess Return (ER) with the ticker SPVXSP which measures the return only from a rolling long position in the first and second month futures contracts.

The total return and excess return versions of the indices are calculated by the following formulas:

$$IndexTR_t = IndexTR_{t-1} \times (1 + CDR_t + TBR_t) \tag{4}$$

$$IndexER_t = IndexER_{t-1} \times (1 + CDR_t), \tag{5}$$

where CDR_t is Contract Daily Return and TBR_t is the Treasury Bill Return earned on the notional value of the position.

The CDR_t and the TBR_t are given by the formulas:

$$CDR_t = \frac{TDWO_t}{TDWI_{t-1}} - 1 \tag{6}$$

$$TBR_t = \left[\frac{1}{1 - \frac{91}{360} \times TBAR_{t-1}}\right]^{\frac{Delta_t}{91}} - 1, \tag{7}$$

where $TDWO_t$ is the Total Dollar Weight Obtained on t and $TDWI_{t-1}$ is the Total Dollar Weight Invested on $t-1$, as determined by the following formulas:

$$TDWO_t = \sum_{i=1}^{2} CRW_{i,t-1} * DCRP_{i,t} \tag{8}$$

$$TDWI_{t-1} = \sum_{i=1}^{2} CRW_{i,t-1} * DCRP_{i,t-1}, \tag{9}$$

where $CRW_{i,t}$ is the Contract Roll Weight of the ith VIX Futures Contract on date t and $DCRP_{i,t}$ is the Daily Contract Reference Price of the ith VIX Futures Contract on date t. $Delta_t$ is the number of calendar days between the current and previous business days and $TBAR_{t-1}$ is the most recent weekly high discount rate for 91-day US Treasury bills effective on the preceding business day.

Inspection of Equations (8) and (9) shows that Contract Daily Return is driven only by the changes in the VIX futures prices and is not dependent on the changes in the weights.

At the start of the roll period, all the weight is allocated to the first month contract. Then, on each subsequent business day a fraction of the first month VIX futures holding is sold and an equal notional amount of the second month VIX futures is bought. The initial position in the first month contract is progressively rolled to the second month futures contract during the course of the month, until the following roll period starts when the old second month VIX futures contract becomes the new first month VIX futures contract. After that the process begins again.

The closing indicative value is linked to the performance of the underlying index less an investor fee. On any calendar date the CIV for the VXX is calculated based on the following equations:

$$CIV_t = CIV_{t-1} \times DIF_t - Fee_t \tag{10}$$

$$DIF_t = \frac{IndexTR_t}{IndexTR_{t-1}} \tag{11}$$

$$Fee_t = \frac{feeRate}{365} \times CIV_t \times DIF_t, \tag{12}$$

where CIV_t is the closing indicative value on any given calendar day t, DIF_t is the daily index factor, $Index_t$ the closing level of the Index, $feeRate$ is the investor fee rate which is equal to 0.89% per year.

Combining the formulas 4, 10, 11 and 12 we arrive at the final formula for the closing indicative value for the given series of ETN:

$$CIV_t = CIV_{t-1} \times (1 + TBR_t + CDR_t) \times (1 - \frac{feeRate}{365}) \tag{13}$$

The closing indicative value for the series of the XIV is equal to:

$$CIV_t = CIV_{t-1} \times DETNP_t - DIF_t \tag{14}$$

$$DETNP_t = 1 + TBR_t + DIP_t \times (-1) \tag{15}$$

$$DIF_t = CIV_{t-1} \times DETNP_t \times \frac{feeRate}{365} \tag{16}$$

$$DIP_t = \frac{IndexER_t}{IndexER_{t-1}} - 1, \tag{17}$$

where $DETNP_t$ is the daily ETN performance, DIF_t is the daily investor fee, DIP_t is the daily index performance, $feeRate$ is equal to 1.35% for the inverse ETN. The daily index performance is adjusted by the leverage amount -1.

Combining Equations (5) and (14)-(17) the closing indicative value for the XIV series is equal to:

$$CIV_t = CIV_{t-1} \times (1 + TBR_t - CDR_t) \times (1 - \frac{feeRate}{365}). \tag{18}$$

Comparing Equations (13) and (18) one can see that though at the first sight the VXX and the XIV are benchmarked to the indices that differ by the return which might be gained through investing in the three-month U.S. Treasury rate, it follows that the closing indicative values of both of the series include it. This observation gives us a ground to assert that two ETNs are cointegrated and are an ideal candidates to analyze their price discovery in Hasbrouck (1995) and Gonzalo and Granger (1995) vector error correction model (VECM) setting.

3 Methodology

Several measures are widely used to quantify the process of price discovery. The most popular are the common factor model (CS) introduced by Gonzalo and Granger (1995) and the information shares (IS) model developed by Hasbrouck (1995). Both of this models are based on the (VECM) between considered securities, though their definition of price discovery differs. The CS measure considers common factor components and the process of error correction to the equilibrium relationship, whereas the IS measure focuses on the contribution of each marker to the variance of the innovation to the common efficient price.

3.1 Error Correction Model

Consider a VECM of the form:

$$\Delta P_t = \alpha(\beta' P_{t-1} - E(\beta' P_{t-1})) + \sum_{i=1}^{k} \Gamma_i \Delta P_{t-i} + e_t \tag{19}$$

where $P_t = (P_{1,t}, P_{2,t})'$ denote a (2×1) vector of (log) prices for the VXX and the XIV respectively, Γ_i is an (2×2) matrix, e_t is a (2×1) vector of the residuals. The VECM includes two parts: the first part, $\alpha(\beta' P_{t-1} - E(\beta' P_{t-1}))$, represents the log-run or equilibrium dynamics between the price series, the second part $\sum_{i=1}^{k} \Gamma_i \Delta P_{t-i}$ depicts the sort-run dynamics induced by market imperfections. Assuming that the cointegrating vector is known, the VECM can be estimated by the ordinary least square regression. The $E(\beta' P_{t-1})$ term captures systematic differences in the prices (such as the difference between a bid and offer quote). This term can be estimated by the sample average prior to the other parameters, corresponding to a "de-meaning" of the data.

3.2 Common Factor Weights

Schwarz and Szakmary (1994) are the first who proposed to use the relative magnitude of the error correction coefficients α_1 and α_2 from Equation (19) to assess the contribution of each market to the price formation process. They argue that the price discovery leader is the market which initiate the mispricing $\beta' P_t$ and the price discovery follower is the market which respond to the disequilibrium. Thus, the price discovery occurs entirely in market 1 if α_1 has zero value or in market 2 if α_2 has zero value.

Gonzalo and Granger (1995) propose the formal derivation for the common factor measure. They use the permanent-transitory (PT) decomposition of P_t in the form

$$P_t = A_1 f_t + A_2 z_t \tag{20}$$

where $f_t = \alpha'_\perp P_t$ is a common factor (the permanent component), $z_t = \beta' P_t$ is the transitory component, $A_1 = \beta_\perp(\alpha'_\perp \beta_\perp)^{-1}$ and $A_2 = \alpha(\beta'\alpha)^{-1}$ are loading matrices. $\alpha_\perp$ and $\beta_\perp$ are (2×1) vectors such that $\beta'_\perp \beta = 0$ and $\alpha'_\perp \alpha = 0$. Making the derivations they impose the condition that the common factor f_t is the linear combination of prices and that the transitory part does not have any permanent effect on the prices. Only the shock to the permanent component can affect the long-run forecast of P_t.

To derive the algebraic expression for the component shares when cointegrating relationship is known we can consider the permanent component of the price $A_1 f_t = \beta_\perp(\alpha'_\perp \beta_\perp)^{-1}\alpha'_\perp P_t \equiv \beta_\perp \gamma' P_t$. Since $\beta = (1, 1)'$, one option for $\beta_\perp$ is $(1, -1)'$. $\alpha_\perp$ can be expressed in terms of the elements of the error correction coefficient vector α. From $\alpha'_\perp \alpha = 0$ follows that $\alpha_{\perp,1}\alpha_1 + \alpha_{\perp,2}\alpha_2 = 0$ which in turn implies that $\alpha_{\perp,1} = -\alpha_{\perp,2}\alpha_2/\alpha_1$. In this way the elements of the vector $\gamma = (\gamma_1, \gamma_2)'$ can be expressend as

$$\gamma_1 = \frac{\alpha_2}{\alpha_1 + \alpha_2}, \gamma_2 = \frac{-\alpha_1}{\alpha_1 + \alpha_2} \tag{21}$$

$$A_1 f_t = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix} \begin{pmatrix} P_{1t} \\ P_{2t} \end{pmatrix} = \begin{pmatrix} \gamma_1 & \gamma_2 \\ -\gamma_1 & -\gamma_2 \end{pmatrix} \begin{pmatrix} P_{1t} \\ P_{2t} \end{pmatrix} \tag{22}$$

The diagonal elements of this matrix represent the component shares (CS) which measure prices discovery between P_1 and P_2. As expected they sum up to one.

3.3 Information Shares

Hasbrouck (1995) defines the price discovery process in terms of the variance of the innovations to the common efficient price. His information shares (IS) measure each market's relative contribution to the variance. If innovations are contemporaneously correlated then IS cannot be computed uniquely and is dependent on the order of the individual asset's price in the price vector. Hasbrouck suggests to use different orders of the prices and compute upper and lower bounds of IS.

Baillie et al. (2002) support the use of the mean of the bounds as an ultimate measure of the market's contribution to the price discovery process.

Let $P_t = (p_{1,t}, ..., p_{n,t})'$ denote an $n \times n$ vector of $I(1)$ log prices. It is assumed that there is a common stochastic component or the efficient price that is shared by all prices. That means that there are $n-1$ cointegrating vectors β_i such that $\beta_i' p_t \sim I(0)$. Even though the prices are nonstationary the difference between any two is stationary and it is convenient to use the following basis of rank $n-1$ for the space of cointegrating vectors:

$$\beta'_{(n-1)\times n} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}. \tag{23}$$

Since ΔP_t is $I(0)$, it has a moving average representation:

$$\Delta P_t = \Psi(L)e_t = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + ..., \tag{24}$$

where $\Psi(L)$ is a matrix polynomial in the lag operator, $\Psi_0 = I_n$, $e_t = (e_{1,t}, ..., e_{n,t})'$ and $e_t \sim iid(0, \Sigma)$ is a zero-mean vector of serially uncorrelated disturbances with covariance matrix Σ. Using the Beveridge-Nelson decomposition we can present the price levels as:

$$P_t = P_0 + \Psi(1)\sum_{i=0}^{t} e_i + \Psi^*(L)e_t \tag{25}$$

where P_0 is a constant $(n \times 1)$ vector of initial values, $\Psi(1)$ is the sum of the moving average coefficients and $\Psi^*(L) = \sum_{k=0}^{\infty} \Psi_k^*$ and $\Psi_k^* = -\sum_{j=k+1}^{\infty} \Psi_j^*$.

The requirements that $\beta' p_t$ is stationary implies that $\beta'\Psi(1) = 0$. Given the structure of β it implies that all of the rows of β are identical and can be expressed as

$$\Psi(1) = 1_n \psi' = \begin{bmatrix} \psi_1 & \cdots & \psi_n \\ \vdots & \ddots & \vdots \\ \psi_1 & \cdots & \psi_n \end{bmatrix}, \tag{26}$$

where $\psi = (\psi_1, ..., \psi_n)$ is an $n \times 1$ vector.

Combining (25) and (26) the price in levels can be written as

$$P_t = P_0 + 1_n \sum_{i=0}^{t} \eta_i^P + \widetilde{\varepsilon}_t, \tag{27}$$

where $\widetilde{\varepsilon}_t = \Psi^*(L)e_t$ is an $I(0)$ pricing error vector, $\eta_t^P = \psi' e_t = \sum_{k=1}^{n} \psi_k e_{kt}$ and $\sum_{i=0}^{t} \eta_i^P$ is the random-walk component that is common to all prices. η_t^P is a permanent shock which is defined as a weighted average of individual market innovations. To interpret the result easier we can write the random-walk component in the following fashion:

$$P_t = P_0 + 1_n m_t + \widetilde{\varepsilon}_t, \tag{28}$$

where $m_t = m_{t-1} + \eta_t^P$. The increment η_i^P is the component of the price change that is permanently impounded into the asset price and is presumably due to new information. The variance of this term is $\psi' \Sigma \psi$. The calculation of IS measure can be divided into two cases. If Σ is diagonal (i.e. the market innovations are not contemporaneously correlated), then the variance consists of the sum of n terms and each of them represents the contribution from a particular market to the innovation to the random walk component of the price.

$$IS_i = \frac{(\psi_i \sigma_i)^2}{\psi' \Sigma \psi}, i = 1, ..., n. \tag{29}$$

If Σ is non-diagonal there is no unique value for the IS, but triangularization approach of the covariance matrix can help to determine lower and upper bounds of IS. In this case the formula is as follows:

$$IS_i = \frac{(\psi' F)_i^2}{\psi' \Sigma \psi}, i = 1, ..., n. \tag{30}$$

Here F is the Cholesky factorization of Σ such that $FF' = \Sigma$ and F is a lower triangular matrix. $(\psi' F)_i$ is the i-th element of $\psi' F$.

In practice, IS is computed from estimates of an vector error correction model of the form 19. The matrix $\Psi(1)$ can be computed using Johansen's factarization and the estimated coefficients by the following formula:

$$\Psi(1) = \beta_{\perp}(\alpha'_{\perp}\Gamma(1)\beta_{\perp})^{-1}\alpha'_{\perp}, \tag{31}$$

where $\beta_{\perp}$ and $\alpha_{\perp}$ are orthogonal vectors satisfying $\beta'\beta_{\perp} = 0$ and $\alpha'\alpha_{\perp} = 0$, respectively. $\Gamma(1) = I_n - \sum_{i=1}^{k-1} \Gamma_i$.

4 Data and Summary Statistics

In this section, we report the daily trading volume of the ETNs, compare the daily prices and describe the properties of the intraday data on the VXX and XIV.

We obtain intraday trade and quote data for the VXX and XIV from the Thomson Reuters Tick History (TRTH) database with millisecond precision. The starting date of the sample is the 3^{th} of January 2011 and the ending date is 31^{st} December 2015. We do not start the sample from the inception of the XIV to avoid any liquidity issues. Tick-by-tick raw data contains different types of data errors. We follow the procedure by Barndorff-Nielsen et al. (2008) to eliminate these errors. During the cleaning procedure we merge the data which correspond to the same second, keeping only the last observation in the group with the same second time stamp.

Table 1 provides the summary statistics on the computed mid point of the bid and ask quotes. Panel A demonstrates the summary statistics for the levels of the VXX and the XIV. The VXX has a positive skewness and displays excess kurtosis. The first-order autocorrelation of the VXX is close to one and the ADF test fails to reject the presence of a unit root at conventional significance levels. The XIV has positive skewness and negative excess kurtosis. The data are highly persistent and an Augmented Dickey-Fuller (ADF) test fails to reject the presence of the unit root. The log transformation which is presented in Panel B smooths the data. The log of the VXX preserves the positive skewness but now exhibit negative kurtosis and the ADF test again demonstrates that the null hypothesis of a unit root can not be rejected. The logarithm values of the levels for the XIV has negative skewness and negative excess kurtosis. According to the ADF test statistic it is not possible to reject the null of a unit root at conventional significance levels showing that these series are non-stationary. Panel C shows the descriptive statistics for the first difference of the logarithmic values of the VXX and the XIV. The ADF tests reject the null hypothesis for a unit root. Therefore, the logarithm of the VXX ans the XIV are $I(1)$ processes.

Figure 1 shows the 5-day moving average of the daily volume of the VXX since the inception to the end of 2015. After one year from the inception, trading volume started to grow and in December 2015 reached about 70 million shares a day showing that this ETN gained acceptance as volatility trading vehicle. The peak average volume occurs in times of market instability. These days are August 2011, when the market event termed as "Black Monday" occured as a response to the USA's credit rating downgrade; May 2012 when the event known as "Grexit" during the European debt crisis took place; June 2013 with the Chinese banking liquidity crisis; October 2014 turmoil sparked by weak US economic data and fears over the health of the global economy; August 2015 market selloff triggered by China's stock market crash.

Figure 2 shows the 5-day moving average daily volume of the XIV from the inception to the end of 2015. The XIV also gained significant growth within one year, though in a lesser scale. By the end of 2015 average daily volume reached 30 million shares, almost twice less than the volume of the VXX. Similar to the VXX, during the big geopolitical events the trading volume in the XIV tends to increase indicating that the traders might change their outlook on the future direction of volatility and take either long or short position in volatility.

The plot of the price series is depicted on Figure 3 which are calculated as the logarithm of mid-quote of the VXX and of the VIX by the end of the trading session. The plot shows that the time series of the prices looks very similar and seem to be a mirror reflection of each other. To facilitate the visualization we present a scaled graph via adding a constant to the negative of the XIV to align the price levels. In Figure 4 it is more notable that the prices replicate each other very closely and that in fact they don't move exactly in parallel, but the gap between them first narrows and start to widen since the end of 2014. We attribute this observation to the slightly different investor fee rates, the influence of which accumulates over time.

5 Empirical results

5.1 Cointegration Test

Before estimating the VECM in Equation (19) we need to test that the studied price series are cointegrated. The ADF tests show that the price series of both the XIV and the VXX are nonstationary. Though they are nonstationary we expect that they do not drift far apart from each other because they have similar underlying indices with the difference that the performance of the XIV is reversed to to the underlying index. Thus, we expect the cointegrating vector to be close to $\beta = (1, 1)'$. We use the Johansen likelihood-ratio test statistic, called the trace statistic, that there is at most one cointegrating vector. Table 2 presents the results of the Johansen procedure applied for each day in the sample. The null hypothesis that the number of cointegrating vectors is less than or equal to one cannot be rejected at the one percent level for 1187 days out of 1257. For the rest of the 70 days the ADF test shows that in the 59 cases one of the price series is stationary (in this case cointegration is not possible). The average cointegrating equation is $(1, 1.002)'$ and the 95% confidence interval for the population mean of the second element of the cointegrating equation is $(0.999, 1.006)$. Thus the null hypothesis that the CE is equal to $(1, 1)'$ cannot be rejected at the 5% level.

5.2 VECM and Daily Price Discovery Measures

The Johansen procedure applied to daily data shows that the cointegrating vector is close to $\beta = (1,1)'$ and the null hypothesis that the population mean is equal to one cannot be rejected at the five percent level. Thus, apart from the exact cointegrating relationship we can take the cointegrating vector $\beta = (1,1)'$ for further investigation.

To identify the lag length of the VECM which will be later estimated for every day in the sample we use the multivariate version of Schwartz's Bayesian Information Criterion (SBIC). After calculating the optimal lag length for every day in the sample we calculate the average value which is equal to 9. Thus to compute our daily measures of price discovery, we consider a VECM model with $\beta = (1,1)'$ and a lag-length $k = 9$. The other three alternatives are $\beta = (1,1)'$ and the original lag-length returned by SBIC; original β returned by the Johansen procedure and $k = 9$; original cointegrating relationship β and VECM lag length k.

Based on the results of the daily estimates of the VECMs for the four different specifications discussed above we calculate the CS using formulas 21 and 22. While calculating these measures for some days there happened to be not as expected results. In the specification with $\beta = (1,1)'$ and $k = 9$, 255 days have the CS which lie outside the boundaries of zero and one, as a consequence of α_1 and α_2 being of different signs. The test for the population mean being equal to one is rejected with $p - value = 0.0002$ for these days. This result might be driven by the deviation of the market prices from the intraday indicative values which approximate the economic value of the series. Even though the ETNs track similar idices the deviation between prices and closing indicative values might cause the deviation from the theoretical cointegrating vector $\beta = (1,1)'$. However the model with the exact cointegrating relationship which was returned by the Johansen procedure and the average optimal lag-length of the daily VECMs provides a better fit and returns only 62 days with CS being outside the interval of zero and one.

Further we calculate the upper and lower bounds of the information share using the equation 30 and after that calculate mean value between the lower and upper bounds.

Table 3 shows distributional properties of both of the measures. The results for CS exclude the days when calculated value of CS is not bounded by zero and one. The results show that the mean value of CS for the VXX is higher than for the XIV in both four methods of calculation. The highest value of the VXX's CS is 56 % and the lowest is 53 %. The highest value for the XIV's CS is 47 % and the lowest is 44 %. The price discovery leadership of the VXX is supported by the IS measure. IS measure states that, on average, price discovery contribution of the VXX varies between 54 - 55 % depending on the method of calculation. As all four methods of the VECM specification show similar results later in our analysis we study only one specification (with the lag-length equal to 9 and the cointegrating relationship returned by the Johansen procedure).

Figure 5 shows the 5-day moving averages computed from the VXX's daily CS and IS. The time-series variation has a similar pattern in both of the measures though CS exhibits higher volatility. We can notice that both of the measures exhibit strong persistence. A more detailed look suggests that this persistence closely follow major geopolitical events discussed earlier in Section 4.

5.3 Determinants of Price Discovery

Numerous studies have focused on the determinants of price discovery between the spot and derivatives markets.

Chu et al. (1999) formulate four possible hypothesis about the features and properties of market structure and security design that can encourage informed traders to choose one particular market to transact. They are the leverage hypothesis, the trading cost hypothesis, the uptick rule hypothesis, and the marketwide information hypothesis. Fleming et. al (1996) study the lead-lag relationship between the stock index, options and futures contracts on this index. They conclude that the leading market is the one which has the lowest trading cost because the informed traders seek to earn the highest profit through execution of theirs trading strategies. Chakravarty et al. (2004), using stocks and options data, show that informed trading in the option market has a significant association with the relative effective bid-ask spread, trading volume and stock volatility.

In summary, the main price discovery determinants between the spot and derivatives market are found to be trading costs, market liquidity measures, volatility of security.

Another strand of research study the influence of investor structure on price discovery mechanism. This work is motivated by the idea that institutional traders are the group of well-informed investors and can improve the informational content of prices.

Boehmer and Kelley (2009) study the effect of institutional shareholdings on the informational efficiency of transaction prices. They show that both institutional holdings and institutional trading activity enhance the price efficiency. Bohl et al. (2011) finds that the price discovery contribution of the futures market is higher during the period of increased institutional ownership. Chen et al. (2015) conduct the price discovery analysis between the S&P 500 exhange-traded fund and S&P 500 index derivatives and show that contribution of SPDR to the price discovery is positively related to its increasing institutional ownership. They highlight that AT and HFT measured by the average trade size enhance the price efficiency. This finding is consistent with the results of Carrion (2013) and Brogaard et al. (2014). Anand and Chakravarty (2007) consider the price discovery mechanism in options market with the relation to trades fragmentation among the informed traders. They confirm the "stealth-trading" hypothesis which states that informed traders camouflage their trades and prefer to transact in medium and small size trades. Thus prior research indicates that retail investors and institutions have a different level of sophistication in their response to new information with the implication that institutions are a group of informed traders.

The area of price discovery between volatility products emerged relatively recently and only a few studies examined the possible determinants of the price discovery process within this asset class. Thus in this section we investigate how the price discovery measures of the VXX are affected by the trading costs, by the volume related metrics, level of volatility and the level of institutional ownership.

To measure trading costs we use relative spread which is calculated as the average value of the difference between ask and bid quotes relative to the quote midpoint. The volume related metrics include the total number of trades, the trading volume, the average time between trades, average trade size, all calculated on a daily basis. As a measure of the relative change in the average time between transactions a ratio of the time between trades may be considered. According to the stealth trading hypothesis informed investors prefer to transact more often and in a smaller trades size. Thus we expect to see that the lower the time between trades in the VXX market the higher the price discovery share of the VXX. In a similar vein the smaller the average trade size of the VXX the higher the price discovery share of the VXX.

We include the VIX in our regression analysis. The level of the VIX is an easily observable variable to the investors, which according to the pricing formula should influence our volatility ETNs in opposite directions. During the times of high levels of fear direct ETNs should benefit and increases in value as during these times futures curve swings from contango to backwardation. Thus institutions might be attracted and use the VXX as a portfolio insurance. In times when market is stable the inverse ETN benefits and proportion of institutions in the direct note should decrease leading to the decrease in price efficiency.

The regression equation is given by:

$$PD_t = \beta_0 + \beta_1 Spread_t + \beta_2 Volume_t + \beta_3 VIX_t + e_t, \tag{32}$$

where PD_t is a price discovery share of the VXX at day t, $Spread_t$ is the ratio of effective spreads at day t, $Volume_t$ is the ratio of one of the volume related metrics at day t, VIX_t is the open level of the VIX index at day t. All the ratios are taken as the VXX value divided by the XIV.

Table 4 shows the correlation matrix between the independent and dependent variables. All the possible explanatory variables prove to be significantly correlated at 1 % level.

Figure 6 depicts the scatter plot between the IS and CS of the VXX and the VIX level. Significant clustering of the observations in the upper right part of the graphs shows that the highest price discovery is associated with high levels of the VIX.

The regression results from Equation (32) are presented in Table 6. All four model specifications provide good fit with the lowest Adjusted R-square being equal to 73.21% and all coefficients' estimates are statistically significant. In line with the previous studies the relative trades number and the trades volume is positively related to the information share metric. Aslo, we find a negative relation between the relative trading cost measure and information share of the VXX. This result suggest that informed traders prefer to trade in the more liquid market and with the lowest transaction costs. The level of the VIX is found to have significant positive impact on the price discovery contribution of the VXX. The negative association between the time between trades and the informational leadership of the VXX and negative association between average trade size and informational leadership of the VXX support the idea of stealth trading. The informed trades tend to break up their trades and transact more often.

5.4 Efficiency and Institutional Ownership

Any institutional investment manager, which is defined as an entity or a person that exercises investment discretion over \$100 millions, must file Form 13F for the Securities and Exchange Commission. Form 13F is a quarterly report of the holdings which called the Official List of Section 13(f) Securitis. The Official List primarily includes U.S. exchange-traded stocks, shares of closed-end investment companies, and shares of exchange traded funds. It also includes certain convertible debt securities, equity options and warrants. The VXX and the XIV can be found there with CUSIP number 06742E711 and 22542D795, respectively. We obtain the historical number of shares outstanding and the numbers of shares held by institutional customers from Bloomberg. Because filings are made quarterly, we have only a snapshot of institutional ownership for every quarter and year in our sample.

We provide an insight at the relationship between institutional ownership and efficiency. The holdings filed by institutions in 13F form are measured at the close of trading on the last trading day of the quarter. To find the relationship between institutional ownership and price discovery measures we calculate the average IS taken around the reporting date within a window of one month. In Table 5, we report averages of the price discovery measures divided into three groups based on the 25th and 75th percentile of institutional holdings. As expected, the price discovery measure of each ETNs increases monotonically with increasing institutional ownership in this ETN. Also, the IS of the VXX is increasing with the relative increase of institutional holdings in the VXX.

Figure 7 illustrates the relationship between the level of VIX, institutional holdings and the informational efficiency of the VXX. Two observations are particularly noteworthy. It is noticeable what the time series graph of both of the price discovery measures of the VXX and the VIX resemble each other. The resemblance intensifies at times when the VIX spikes. The infrequent quarterly numbers do not allow to make an exact correspondence between the increased institutional holdings and the increased informational efficiency but the general pattern is that at times of high levels of the VIX the institutions tend to hold the VXX and don't invest heavily in the XIV. At times when the market does not experience much volatility, institutions reduce their holdings in the VXX and invest more in the XIV. Subsequently that leads to the more efficient pricing of the VXX during the market instability and to the more efficient pricing of the XIV during relatively stable market conditions as the institutions switch between the VXX and the XIV depending on the market conditions. As one particular example, in August 2011, a month that experienced significant market volatility due to what was termed as "Black Monday" which happened on August 8, the daily volume of the VXX reached over 140 mln contracts. The level of VIX closed at 48 that day and stayed high during the next four months. Amid this turmoil institutions were trying to protect their portfolios and were buying the VXX. Institutional holdings reached 370% and 290% as of the end of August and December 2014 respectively. That period the VXX served as an insurance for their portfolios because the futures curve was in backwardation and did not suffer from contango losses allowing the value of the VXX to grow.

5.5 Order Imbalance

In this section we use daily measures of order imbalance for retail and institutional trades to see if it significantly influence price discovery of our volatility ETNs.

We classify each transaction as a sell or a buy order using a modified algorithm proposed by Lee and Ready (1991). Specifically, we use a quote rule and compare a trade price with a midpoint of the most recent bid-ask quote. If the trade price is above the midpoint, then the trade classified as buyer initiated, if the trade price is below the midpoint, then the trade is classified as seller initiated. In case the midpoint is equal to the trade price, we don't use a tick test, but leave the transaction unclassified. To infer a trade type, individual or institutional, we use dollar-based cutoffs similar to Barber et al. (2009). They use results outlined in Lee and Radhakrishna (2000, Table 6). Trades less than \$5000 are used as a proxy for individual investor trades (small trades), whereas trades greater than \$50,000 are used as a proxy for institutional investor trades (large trades). A "buffer zone" reduces a probability of erroneous inclusion of an institutional trade in small trades and an individual trade in large trades. Lee and Radhakrishna (2000) use a dataset of NYSE stocks during a three-month period in 1990-1991, thus to account for inflation rate we compare the consumer price index in 1991 and 2011 and adjust cut-off points accordingly. The procedure results in \$8220 and \$82200 cut-off points for small and large trades respectively.

Easley et al. (1996) introduce the probability of informed trading (PIN) measure which represents a proportion of informed trades among all trades. Their measure directly depends on the number of buy and sell trades taking place in the market. The occurrence of an order imbalance does not necessarily imply that some group of traders possesses private information but could arise from a random trading event which would result in a temporary price deviation. However, Chen and Choi (2012) use the PIN variable for Canadian shares listed on both Canadian and U.S. stock exchanges and find that a larger information share is associated with a higher PIN measure. Thus their study shows that the order imbalance is associated with private information which is revealed through trading activity and permanently impound in prices.

Puckett and Yan (2011) directly asses the informational advantage of institutional investors by analyzing a broad set of high-frequency institutional trades. They conclude that institutions earn a significant abnormal returns on their trades within a quarter and that such a positive performance is persistent.

To quantify order imbalance for each day t we introduce two measures which we compute separately for two groups of investors $g \in \{retail, institutional\}$ and also for the sum of these two groups.

$$OINum_{gnt} = \frac{BuyNum_{gnt} - SellNum_{gnt}}{BuyNum_{gnt} + SellNum_{gnt}} \tag{33}$$

$$OIVol_{gnt} = \frac{BuyVol_{gnt} - SellVol_{gnt}}{BuyVol_{gnt} + SellVol_{gnt}}, \tag{34}$$

where n is either the VXX or the XIV. The first measure shows the difference between the number of buyer-initiated trades less the number of seller-initiated trades relative to the sum of the number of the buy and sell trades. The second measure instead of the number of trades considers the difference between the buyer-initiated shares purchased less the seller-initiated shares sold relative to the sum of the number of shares bought and sold on day t, and thus is a measure of imbalance of trading volume.

To asses the influence of order imbalance on price discovery measure we split our analysis in two parts. First we run a regression in the form:

$$PD_t = \beta_0 + \beta_1|OI_{VXX}| + \beta_2|OI_{XIV}| + e_t \tag{35}$$

In our first specification, by taking the absolute values, we treat both excessive buying and selling equally. In the second specification we consider the influence of each quartile of order imbalance on price discovery measure via piecewise regression:

$$\begin{aligned} PD_t = \beta_0 &+ \beta_1 D^{0-25}|OI_{VXX}| + \beta_2 D^{25-50}|OI_{VXX}| + \beta_3 D^{50-75}|OI_{VXX}| \\ &+ \beta_4 D^{75-100}|OI_{VXX}| + \beta_5 D^{0-25}|OI_{XIV}| + \beta_6 D^{25-50}|OI_{XIV}| \\ &+ \beta_7 D^{50-75}|OI_{XIV}| + \beta_8 D^{75-100}|OI_{XIV}| + e_t, \end{aligned} \tag{36}$$

where

$$D^{i-j} = \begin{cases} 1, & \text{if } i^{th} \text{ percentile} \leq \text{order imbalance} \leq j^{th} \text{ percentile} \\ 0, & \text{otherwise,} \end{cases}$$

$i, j \in \{0\%, 25\%, 50\%, 75\%, 100\%\}$.

The results of both of these specifications are presented in Tables 7 and 8. Table 7 considers order imbalance of the number of trades, Table 8 of the number of shares (Equations (33-34) respectively). Panel A of both of the tables suggests that order imbalance in the VXX decreases price discovery while order imbalance in the XIV improves it. We also can notice a disparate impact on price discovery between institutional and retail investors with the effect of order imbalance among retail investors being stronger. Panel B shows the results of a piecewise regression. Now we can see that only excessive selling or buying (captured by dummies D^{0-25} and D^{75-100}) has a significant impact of price discovery. The magnitude of coefficients has different values thus suggesting that there might be asymmetry between ask-side and buy-side orders.

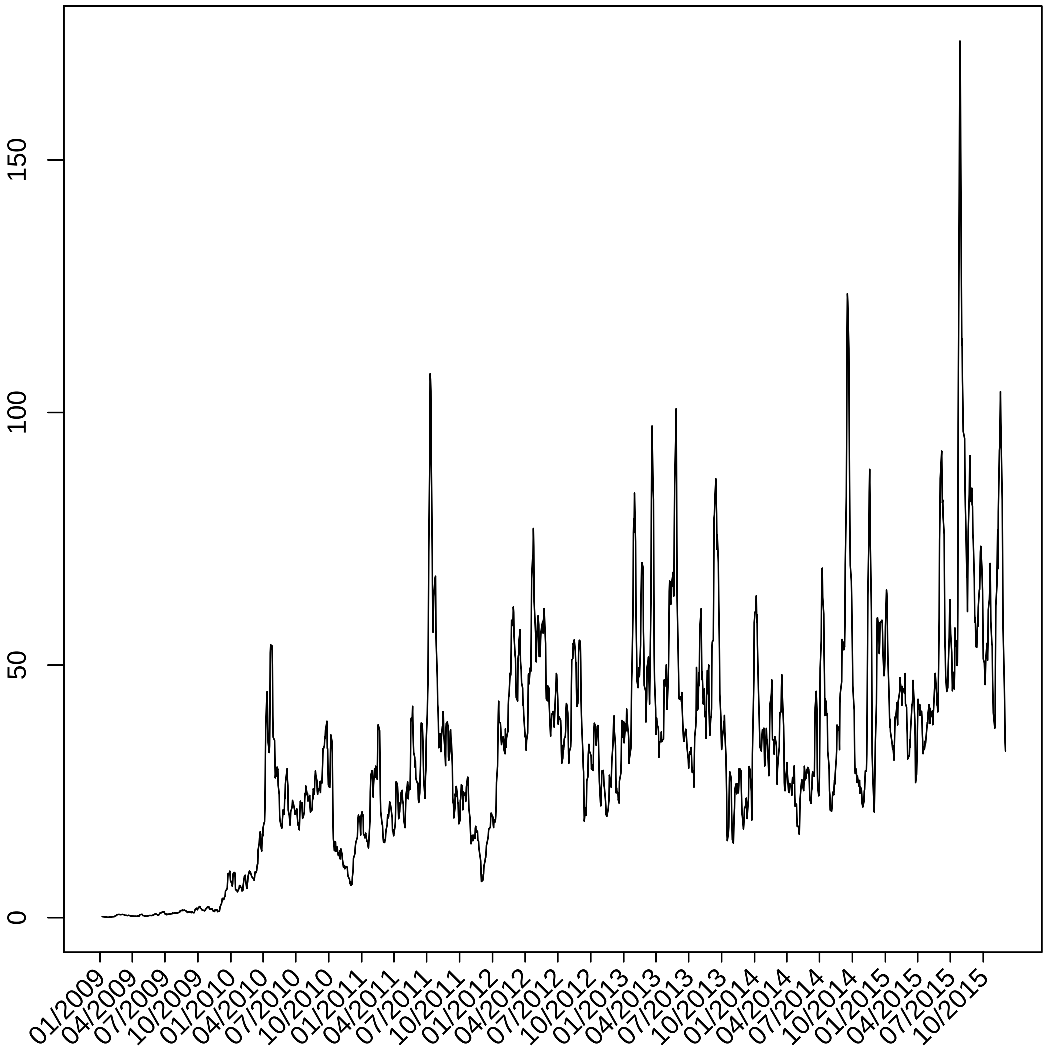
One possible explanation for such an effect might be that noisy signals of retail traders expressed as either positive or negative order imbalance in the XIV improves capacity of the price of the VXX to impound new information and at the same time noisy signals in the VXX trading deteriorate the efficiency of the VXX. This order imbalance might be caused by the overreaction of retail investors. At the same time institutional investors might be less influenced by market sentiments and their order imbalances have lesser impact. These reasonings are in line with the findings of Chang et al. (2013) which report that order imbalance of both individual and institutional investors contribute to asset price formation and that trading behavior of individual investors is strongly affected by market conditions. Additionally they find that order imbalance captured by volume of trades is affected higher than the one captured by the number of trades.

6 Conclusions

In this paper the examine the intraday price discovery relation between the VIX short-term futures ETN and inverse VIX short term futures ETN. To use the approaches of Hasbrouck (1995) and Gonzalo and Granger (1995) we first conduct the unit root test, cointegration analysis and build the VECM. The Johansen approach suggest that the two series are cointegrated, thus a long-run equilibrium relationship exist. We find that both information share and component share show the similar time variation in the price discovery measures with the CS being more volatile. We conduct a times series regression between the daily price discovery measures and several possible determinants. We find that the price discovery process is associated with the relative trading costs, with the relative trading volume and number of trades. VXX prices tend to be more informative on average when the relative trading volume in the VXX is high and when the XIV volume low, when the effective bid-ask spread in the VXX narrows and in the XIV widens. We find that when the trade size and the time between trades of the VXX decrease relative to the XIV the contribution to price discovery made by the VXX increases. We further document that when expected future market volatility increases the traders are more likely to transact in the VXX market. We also provide evidence that during the high institutional ownership in either of the products the price discovery contribution of that product increases. This analysis suggests that the institutions' investment strategy is based on the current market conditions and that they tend to buy the VXX during the market turmoils and invest more in the XIV when there are no significant market volatility.

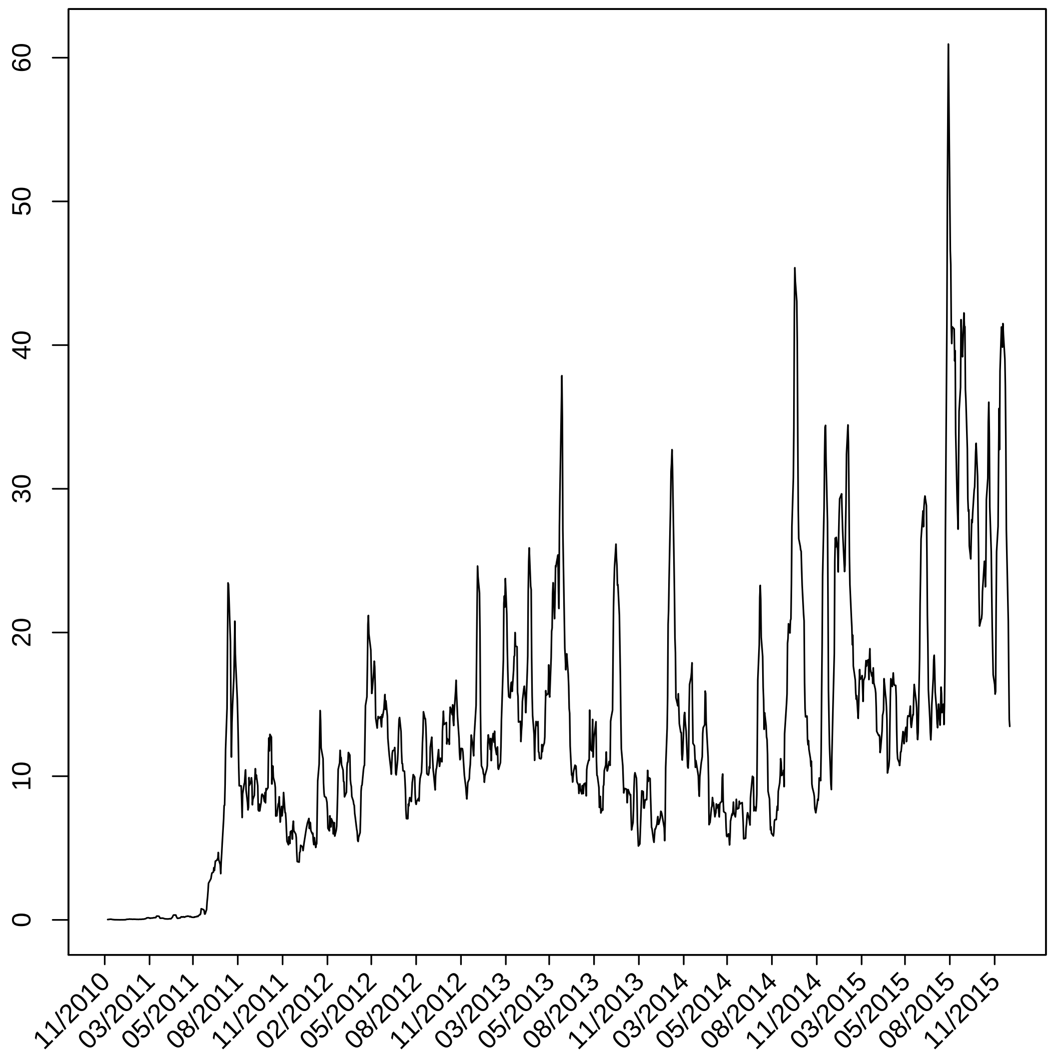
However, when we consider daily order imbalances among institutional trades we cannot provide a proof of a significant effect of buying pressure of institutions on price discovery and their stabilizing role. We find the same direction of influence of order imbalances for both retail and institutional investors with retail trades have higher impact. We document that order imbalance impairs price discovery function and thus can be interpreted as a measure of noise presumably due to the effect of market sentiments.

Figure 1: Daily volume for the VXX (in 1,000,000s)



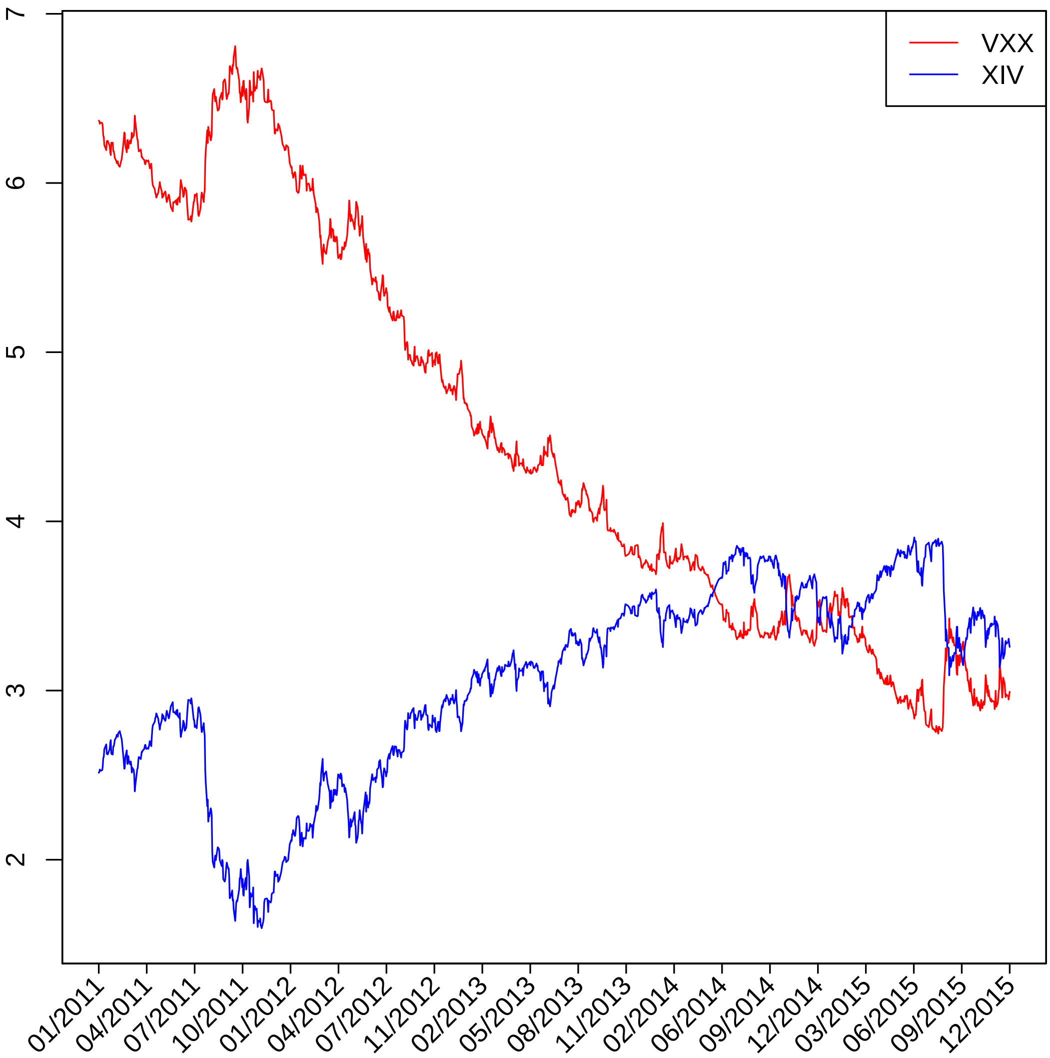
Note: This figure shows the daily trading volume for the VXX for the period between January 29, 2009 and December 31, 2015.

Figure 2: Daily volume for the XIV (in 1,000,000s)



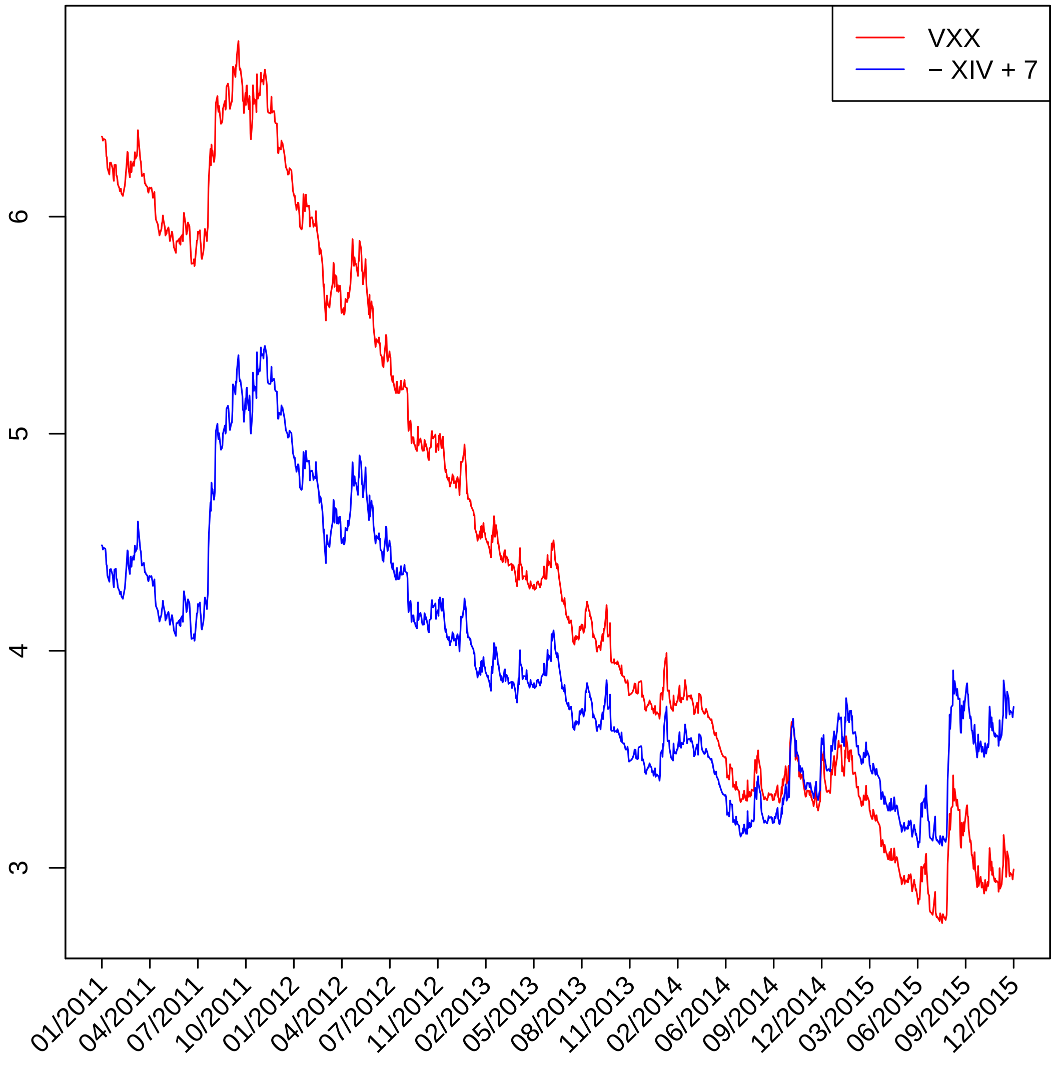
Note: This figure shows the daily trading volume for the XIV for the period between November 30, 2010 and December 31, 2015.

Figure 3: Daily closing prices for the VXX and the XIV



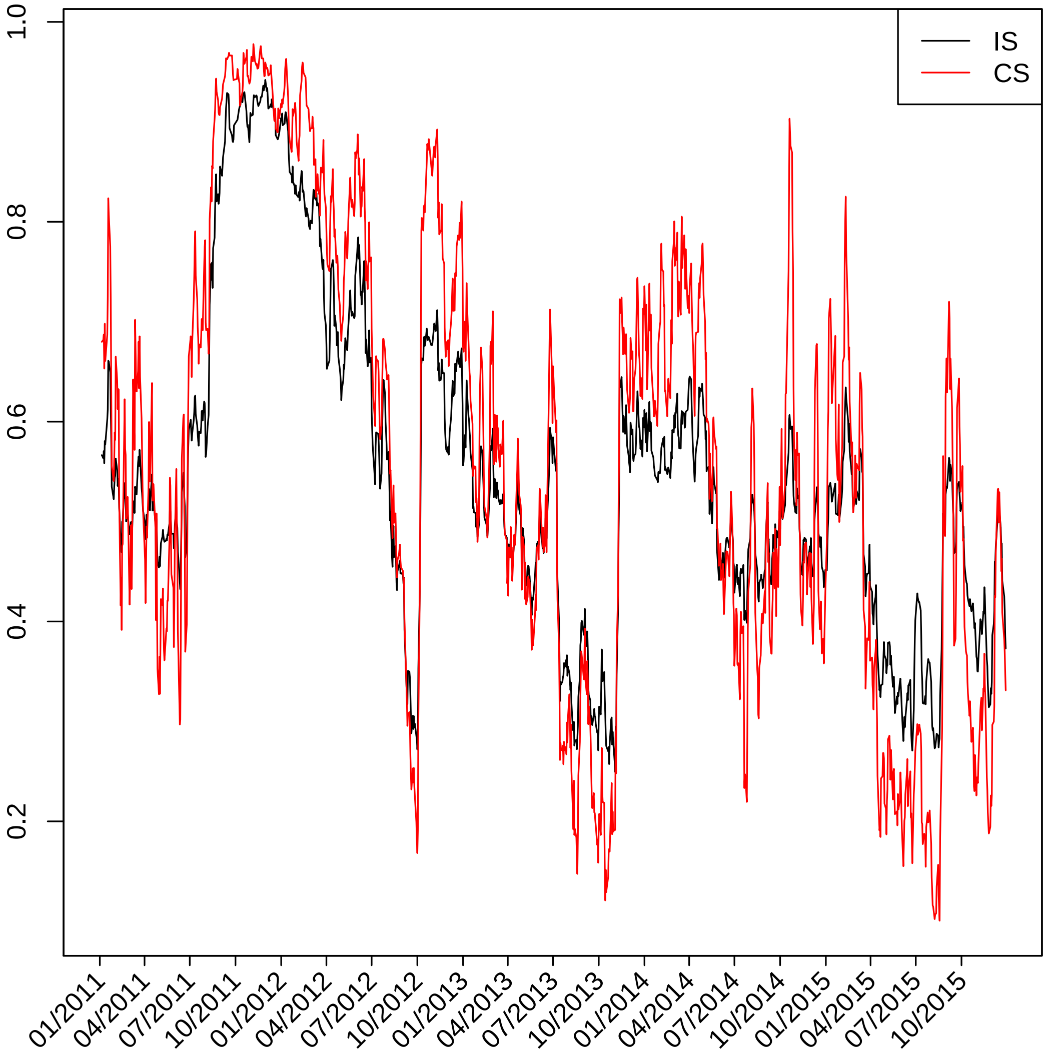
Note: This figure shows a daily log price history for the VXX and XIV from January 3, 2011 to December 31, 2015.

Figure 4: Scaled daily closing prices for the VXX and the -XIV



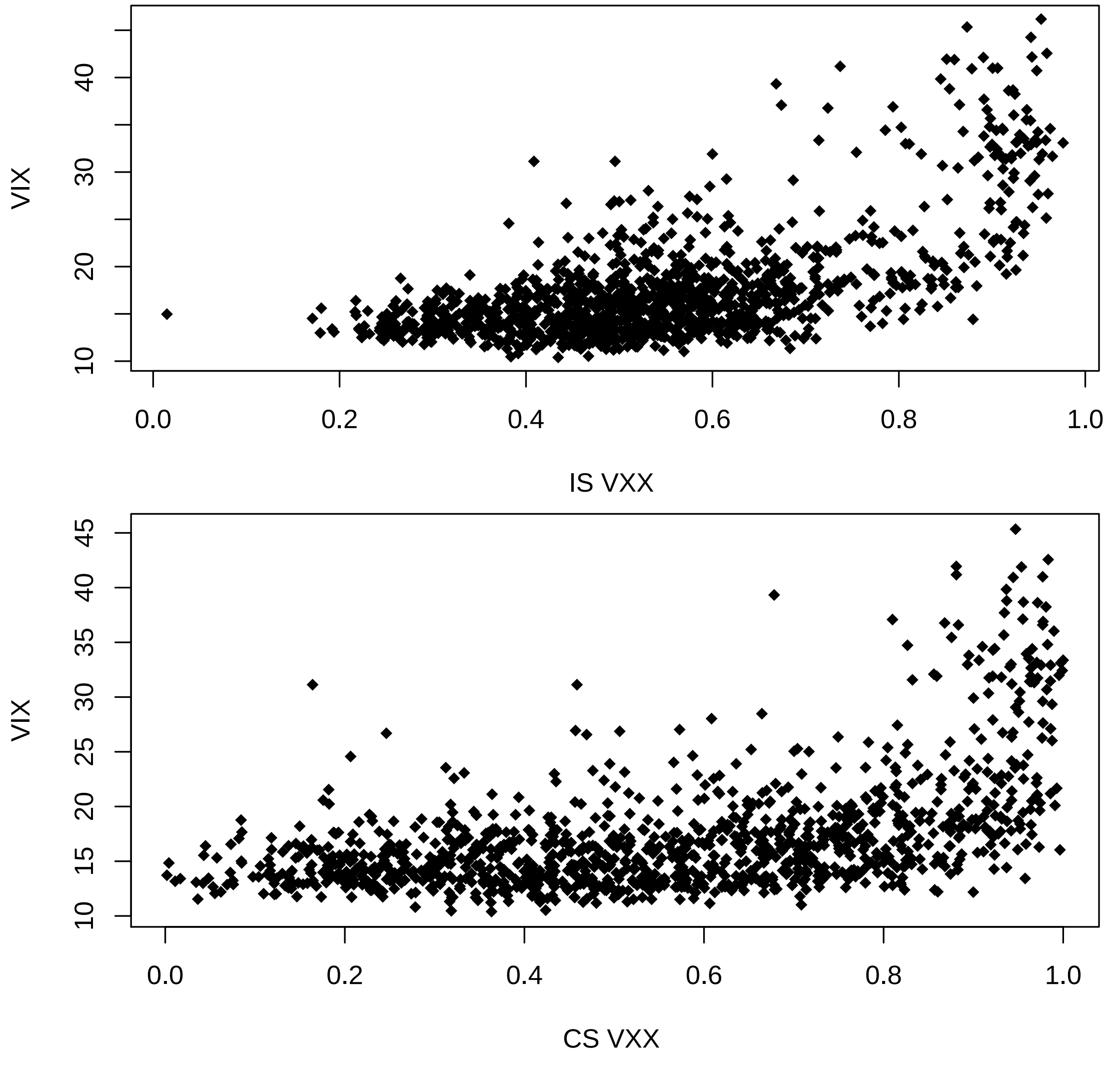
Note: This figure shows a daily log price history for the VXX and XIV indexed to similar values for side-by-side comparison from January 3, 2011 to December 31, 2015.

Figure 5: Price discovery measures for the VXX



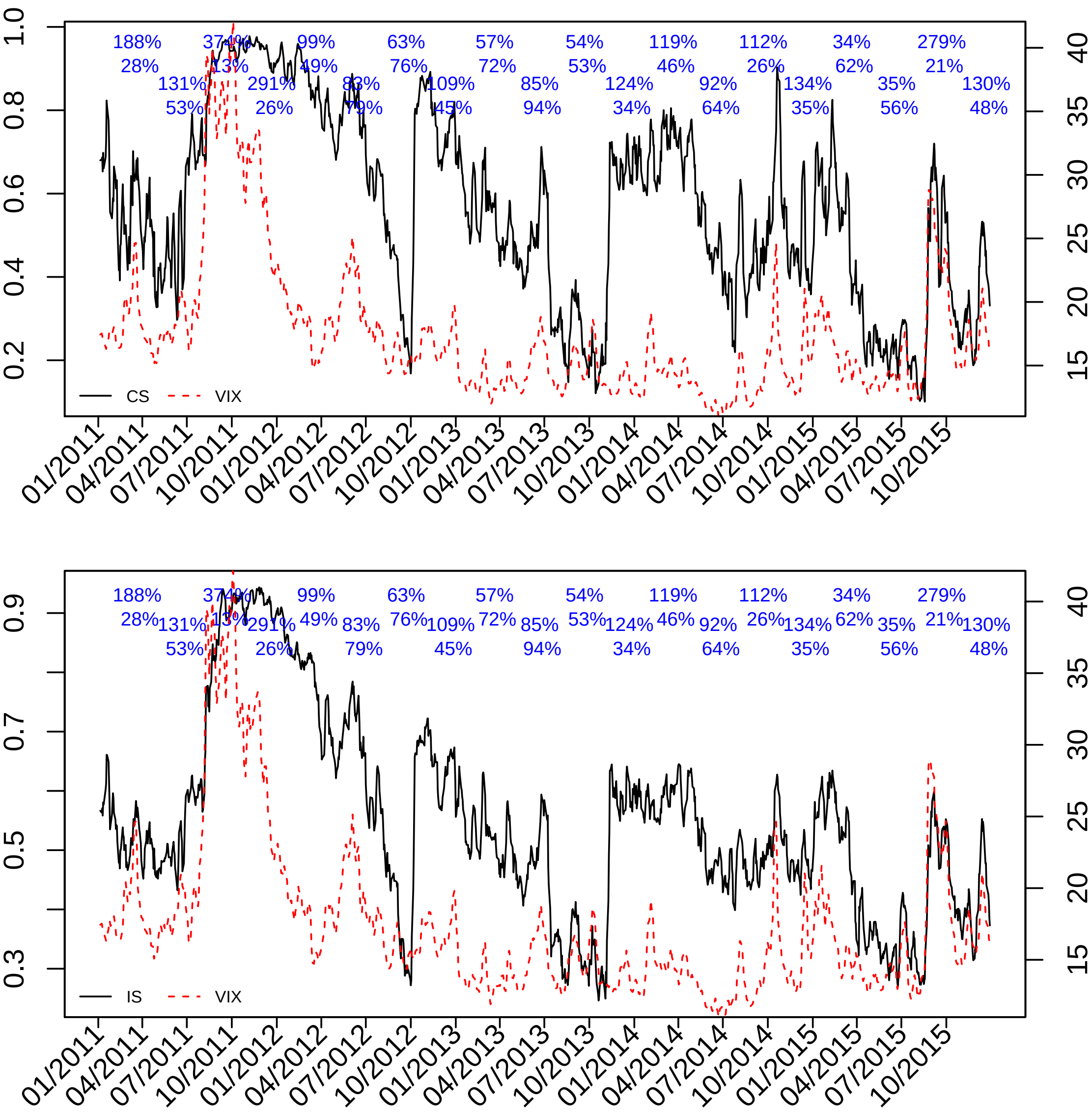
Note: This figure depicts the evolution of price discovery measures for the VXX between January 3, 2011 and December 31, 2015.

Figure 6: Price discovery measures of the VXX with relation to the opening level of the VIX index



Note: This figure depicts the scatter plot between the daily IS (upper panel) and CS (lower panel) measures and the opening level of the VIX.

Figure 7: Price discovery measures for the VXX with relation to the level of VIX and quarterly Institutional holdings



Note: This figure depicts the time series of the Price discovery measures and the level of VIX index. The numbers in blue show the the percentage of institutional holding in the VXX (top) and in the XIV (bottom).

Table 1: Descriptive Statistics

	VXX	XIV
	Panel A: Levels	
Mean	183.59	24.26
St. Dev.	211.09	11.79
Skewness	1.29	0.22
Kurtosis	0.53	-0.97
$\rho(1)$	0.9999	0.9999
ADF	-1.78**	-1.65**
JB Stat $\times 10^6$	8.31	1.34
	Panel B: Log	
Mean	4.51	3.04
St. Dev.	1.21	0.58
Skewness	0.32	-0.63
Kurtosis	-1.33	-0.51
$\rho(1)$	0.9999	0.9999
ADF	-0.92***	-1.41***
JB Stat $\times 10^6$	2.61	2.22
	Panel C: Log Differences	
Mean	0	0
St. Dev.	0	0
Skewness	63.62	-87.11
Kurtosis	82285.34	100326.7
$\rho(1)$	-0.022	-0.035
ADF	-3854.84	-3898.5
JB Stat $\times 10^{15}$	8.09	1.2

Note: This table presents descriptive statistics on the mid quote of the VXX and the XIV at 1-second frequency for the full sample. $\rho(1)$ denotes the first-order autocorrelation; The JB statistic is Jarque-Bera statistic. the augmented Dickey-Fuller (ADF) statistics test the null hypothesis that an examined series has a unit root. **, *** indicates significance at the 5% and 1% levels, respectively.

Table 2: Cointegrating equation estimation results

No. of days	No. of days when indicated one CE	mean CE	95% CI
1257	1187*** 1077**	(1, 1.002)	(0.999, 1.006)

Note: This table reports the results of the Johansen trace statistics tests on the VXX and the XIV for each day in the sample. **, *** denotes significance at the 5% and 1% levels, respectively.

Table 3: Price Discovery Measures Descriptive Statistics

	VXX		XIV	
	IS	CS	IS	CS
Panel A: Exact Avg				
Mean	0.549	0.561	0.451	0.439
5%	0.291	0.153	0.095	0.049
Median	0.529	0.570	0.471	0.430
95%	0.905	0.951	0.709	0.847
Std. Dev.	0.174	0.251	0.174	0.251
Skewness	0.462	-0.100	-0.462	0.100
Kurtosis	-0.159	-1.009	-0.159	-1.009
JB Stat	46.017	52.358	46.017	52.358
No. excluded days	-	62	-	62
Panel B: Exact Exact				
Mean	0.549	0.561	0.451	0.439
5%	0.290	0.141	0.096	0.044
Median	0.533	0.572	0.467	0.428
95%	0.904	0.956	0.710	0.859
Std. Dev.	0.173	0.254	0.173	0.254
Skewness	0.450	-0.105	-0.450	0.105
Kurtosis	-0.190	-1.006	-0.190	-1.006
JB Stat	44.294	52.325	44.294	52.325
No. excluded days	-	61	-	61
Panel C: (1,1) Exact				
Mean	0.543	0.530	0.457	0.470
5%	0.242	0.090	0.097	0.046
Median	0.529	0.530	0.471	0.470
95%	0.903	0.954	0.758	0.910
Std. Dev.	0.198	0.271	0.198	0.271
Skewness	0.181	-0.046	-0.181	0.046
Kurtosis	-0.703	-1.105	-0.703	-1.105
JB Stat	32.442	50.243	32.442	50.243
No. excluded days	-	269	-	269
Panel D: (1,1) Avg				
Mean	0.543	0.529	0.457	0.471
5%	0.234	0.087	0.093	0.042
Median	0.533	0.531	0.467	0.469
95%	0.907	0.958	0.766	0.913
Std. Dev.	0.197	0.270	0.197	0.270
Skewness	0.195	-0.036	-0.195	0.036
Kurtosis	-0.640	-1.097	-0.640	-1.097
JB Stat	29.119	50.118	29.199	50.118
No. excluded days	-	255	-	255

Table 4: Correlation Matrix: Determinants of Price Discovery Measure

	IS	Relative Spread	Trades Num	Trades Volume	Time b/w Trades	Avg Trades Volume	VIX
IS	1.000	-0.857	0.774	0.097	-0.804	-0.841	0.604
Relative Spread	-0.857***	1.000	-0.627	-0.015	0.836	0.930	-0.530
Trades Num	0.774***	-0.627	1.000	0.367	-0.767	-0.666	0.628
Trades Volume	0.097***	-0.015	0.367	1.000	-0.393	0.101	-0.084
Time b/w Trades	-0.804***	0.836	-0.767	-0.393	1.000	0.812	-0.446
Avg Trades Volume	-0.841***	0.930	-0.666	0.101	0.812	1.000	-0.546
VIX	0.604***	-0.530	0.628	-0.084	-0.446	-0.546	1.000

Note: This table presents the correlation matrix of the price discovery determinants. IS is the Information Share of the VXX, Relative Spread is the ratio of relative spreads , Trades Num is the ratio of the number of trades, Trades Volume is the ratio of the trading volumes, Avg Trades Volume is the ratio of the average trading volumes, Time b/w Trades is the ration of the average time between trades. All ratios are taken as the VXX per the XIV on the daily basis. *** is used to indicate significance at the 1% level.

Table 5: Institutional effect on efficiency

	Intitutional holidings		
	<25th percentile	25th-75th percentile	>75th percentile
Absolute IS VXX	47.1%	53.8%	63.1%
Absolute IS XIV	37.8%	46.9%	52.4%
Relative IS VXX to XIV	44.4%	55.8%	62.2%

Note: This table shows the effect of institutional holdings on efficiency. We form three groups of the institutional holdings (IH) quantiles and calculate the mean value of the IS within one month window around the reporting date. First two raws show the IS in absolute terms, the third raw presents the dependence of the VXX's IS on the relative IH of the VXX to the XIV.

Table 6: Determinants of IS

Intercept	0.5969***	0.5737***	0.6680***	0.6728***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Relative Spread	-0.1633***	-0.2148***	-0.1471***	-0.1509***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Trades Num	0.0232***	-	-	-
	(0.0000)			
Trades Volume	-	0.0166***	-	-
		(0.0000)		
Time b/w Trades	-	-	-0.2572***	-
			(0.000)	
Avg Trade Vol	-	-	-	-0.0783***
				(0.0000)
VIX	0.0020***	0.0087***	0.0081***	0.0074***
	(0.0092)	(0.0000)	(0.0000)	(0.0000)
Adjusted R^2 (%)	73.21	77.55	79.00	77.33

Note: This table reports the coefficient estimates of regression given by equation 32. P-values are reported in parenthesis and ***, ** and * are used to indicate significance at the 1%, 5% and 10% respectively.

Table 7: Effect of OI measured by number of trades on price discovery

Panel A: Regression on absolute values of OI			
	Institutional	Retail	Both
Intercept	0.5328***	0.5416***	0.5793***
	(0.0000)	(0.0000)	(0.0000)
$\lvert OI.VXX \rvert$	-0.3874**	-0.8986***	-0.8428***
	(0.0282)	(0.0000)	(0.0000)
$\lvert OI.XIV \rvert$	0.4947***	1.4064***	0.2600***
	(0.0000)	(0.0000)	(0.0003)
Adjusted R^2 (%)	5.20	9.37	2.78
Panel B: Piecewise regression on absolute values of OI			
Intercept	0.5159***	0.5355***	0.6155***
	(0.0000)	(0.0000)	(0.0000)
$\lvert OI.VXX^{0-25} \rvert$	-0.4113*	-1.2604***	-1.6339***
	(0.0838)	(0.0000)	(0.0000)
$\lvert OI.VXX^{25-50} \rvert$	-0.0405	4.0044**	-0.5214
	(0.9562)	(0.0323)	(0.7856)
$\lvert OI.VXX^{50-75} \rvert$	1.7829	0.3091	-1.5638*
	(0.2864)	(0.7052)	(0.0776)
$\lvert OI.VXX^{75-100} \rvert$	-0.1653	-0.4128**	-0.5393***
	(0.4636)	(0.0380)	(0.0095)
$\lvert OI.XIV^{0-25} \rvert$	0.5284***	1.010***	0.0393
	(0.0000)	(0.0000)	(0.6782)
$\lvert OI.XIV^{25-50} \rvert$	0.6026	0.3770	-4.7394***
	(0.4037)	(0.8035)	(0.0002)
$\lvert OI.XIV^{50-75} \rvert$	1.1082**	-0.0865	-1.7390**
	(0.0286)	(0.9208)	(0.0213)
$\lvert OI.XIV^{75-100} \rvert$	0.4962***	1.4362***	0.2698***
	(0.0000)	(0.0000)	(0.0064)
Adjusted R^2 (%)	5.09	11.05	5.39

Note: This table shows the effect of order imbalance of institutional, retail and both types of investors on price discovery measure. The order imbalance is measured as the difference between the number of buy and sell trades relative to the overall number of trades.

Table 8: Effect of OI measured by volume of on price discovery

Panel A: Regression on absolute values of OI			
	Institutional	Retail	Both
Intercept	0.5082***	0.5285***	0.5353***
	(0.0000)	(0.0000)	(0.0000)
$\|OI.VXX\|$	-0.0160	-0.8284***	-0.2034
	(0.864)	(0.0000)	(0.1110)
$\|OI.XIV\|$	0.3350***	1.5672***	0.2917***
	(0.0000)	(0.0000)	(0.0000)
Adjusted R^2 (%)	5.54	10.89	3.18
Panel B: Piecewise regression on absolute values of OI			
Intercept	0.5162***	0.5430***	0.5289***
	(0.0000)	(0.0000)	(0.0000)
$\|OI.VXX^{0-25}\|$	0.0549	-1.3629***	-0.0570
	(0.637)	(0.0000)	(0.732)
$\|OI.VXX^{25-50}\|$	-0.0525	1.8648	0.3877
	(0.918)	(0.3661)	(0.613)
$\|OI.VXX^{50-75}\|$	0.7833	-0.4485	1.6563
	(0.426)	(0.5885)	(0.146)
$\|OI.VXX^{75-100}\|$	-0.0841	-0.4172**	-0.1703
	(0.530)	(0.0504)	(0.317)
$\|OI.XIV^{0-25}\|$	0.2975***	1.1980***	0.2540
	(0.0000)	(0.0000)	(0.0000)
$\|OI.XIV^{25-50}\|$	-0.1178	-2.4073	-0.2064
	(0.679)	(0.1267)	(0.621)
$\|OI.XIV^{50-75}\|$	0.0075	0.4368	0.0333
	(0.989)	(0.6328)	(0.962)
$\|OI.XIV^{75-100}\|$	0.3426***	1.5012***	0.3267***
	(0.0000)	(0.0000)	(0.0000)
Adjusted R^2 (%)	5.46	12.11	3.14

Note: This table shows the effect of order imbalance of institutional, retail and both types of investors on price discovery measure. The order imbalance is measured as the difference between the volume of buy and sell trades relative to the overall volume of trades.

References

Alexander, C. and Korovilas, D. (2012*a*), 'Diversification of equity with VIX futures: Personal views and skewness preference', *Available at SSRN 2027580* .

Alexander, C. and Korovilas, D. (2012*b*), 'Understanding ETNs on VIX futures', *Available at SSRN 2043061* .

Anand, A. and Chakravarty, S. (2007), 'Stealth trading in options markets', *Journal of Financial and Quantitative Analysis* **42**(01), 167–187.

Bahaji, H. and Aberkane, S. (2015), 'How rational could VIX investing be?', *Economic Modelling* .

Barber, B. M., Odean, T. and Zhu, N. (2009), 'Do retail trades move markets?', *Review of Financial Studies* **22**(1), 151–186.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2009), 'Realized kernels in practice: Trades and quotes', *The Econometrics Journal* **12**(3), 1–32.

Boehmer, E. and Kelley, E. K. (2009), 'Institutional investors and the informational efficiency of prices', *Review of Financial Studies* **22**(9), 3563–3594.

Bohl, M. T., Salm, C. A. and Schuppli, M. (2011), 'Price discovery and investor structure in stock index futures', *Journal of Futures Markets* **31**(3), 282–306.

Bollen, N. P. and Whaley, R. E. (2004), 'Does net buying pressure affect the shape of implied volatility functions?', *The Journal of Finance* **59**(2), 711–753.

Bordonado, C., Molnár, P. and Samdal, S. R. (2016), 'VIX exchange traded products: Price discovery, hedging, and trading strategy', *Journal of Futures Markets* .

Brogaard, J., Hendershott, T. and Riordan, R. (2014), 'High-frequency trading and price discovery', *Review of Financial Studies* **27**(8), 2267–2306.

Carr, P. and Wu, L. (2006), 'A tale of two indices', *The Journal of Derivatives* **13**(3), 13–29.

Carrion, A. (2013), 'Very fast money: High-frequency trading on the NASDAQ', *Journal of Financial Markets* **16**(4), 680–711.

Chakravarty, S. (2001), 'Stealth-trading: Which traders' trades move stock prices?', *Journal of Financial Economics* **61**(2), 289–307.

Chakravarty, S., Gulen, H. and Mayhew, S. (2004), 'Informed trading in stock and option markets', *The Journal of Finance* **59**(3), 1235–1257.

Chang, M. C., Tsai, C.-L., Wu, R. C.-F. and Zhu, N. (2013), Sentiment, market order choice, and returns.

Chen, H. and Choi, P. M. S. (2012), 'Does information vault niagara falls? cross-listed trading in new york and toronto', *Journal of Empirical Finance* **19**(2), 175–199.

Chen, W. P., Chung, H. and Lien, D. (2015), The effect of institutional ownership on price discovery.

Chu, Q. C., Hsieh, W.-L. G. and Tse, Y. (1999), 'Price discovery on the s&p 500 index markets: An analysis of spot index, index futures, and spdrs', *International Review of Financial Analysis* **8**(1), 21–34.

Daigler, R. T. and Rossi, L. (2006), 'A portfolio of stocks and volatility', *The Journal of Investing* **15**(2), 99–106.

Dash, S. and Moran, M. T. (2005), 'VIX as a companion for hedge fund portfolios', *The Journal of Alternative Investments* **8**(3), 75.

DeLisle, J., Doran, J. and Krieger, K. (2014), 'Volatility as an asset class: Holding vix in a portfolio', *Available at SSRN 2534081* .

Demeterfi, K., Derman, E., Kamal, M. and Zou, J. (1999), 'More than you ever wanted to know about volatility swaps', *Goldman Sachs quantitative strategies research notes* **41**.

Deng, G., McCann, C. J. and Wang, O. (2012), 'Are VIX futures ETPs effective hedges?', *The Journal of Index Investing* **3**(3), 35–48.

Easley, D., Kiefer, N. M., O'hara, M. and Paperman, J. B. (1996), 'Liquidity, information, and infrequently traded stocks', *The Journal of Finance* **51**(4), 1405–1436.

Fleming, J., Ostdiek, B. and Whaley, R. E. (1996), 'Trading costs and the relative rates of price discovery in stock, futures, and option markets', *Journal of Futures Markets* **16**(4), 353–387.

Frijns, B., Tourani-Rad, A. and Webb, R. I. (2015), 'On the intraday relation between the VIX and its futures', *Journal of Futures Markets* .

Gonzalo, J. and Granger, C. (1995), 'Estimation of common long-memory components in cointegrated systems', *Journal of Business & Economic Statistics* **13**(1), 27–35.

Husson, T. and McCann, C. J. (2011), 'The vxx etn and volatility exposure', *PIABA Bar Journal* **18**(2).

Lee, C. M. and Radhakrishna, B. (2000), 'Inferring investor behavior: Evidence from torq data', *Journal of Financial Markets* **3**(2), 83–111.

Lee, C. and Ready, M. J. (1991), 'Inferring trade direction from intraday data', *The Journal of Finance* **46**(2), 733–746.

Puckett, A. and Yan, X. S. (2011), 'The interim trading skills of institutional investors', *The Journal of Finance* **66**(2), 601–633.

Schwarz, T. V. and Szakmary, A. C. (1994), ‘Price discovery in petroleum markets: arbitrage, cointegration, and the time interval of analysis’, *Journal of Futures Markets* **14**(2), 147–167.

Shu, J. and Zhang, J. E. (2012), ‘Causality in the VIX futures market’, *Journal of Futures Markets* **32**(1), 24–46.

Signori, O., Briere, M. and Burgues, A. (2010), ‘Volatility exposure for strategic asset allocation’, *Journal of Portfolio Management* **36**(3), 105–116.

Whaley, R. E. (2009), ‘Understanding VIX’, *The Journal of Portfolio Management* **26**(3), 12–17.

Whaley, R. E. (2013), ‘Trading volatility: At what cost?’, *Journal of Portfolio Management* **40**(1), 95.

Whaley, R. E., O’Neill, M. J. and Bollen, N. P. (2014), On the supply of and demand for volatility. Working paper.