An Impact of the Introduction of ThaiDEX SET50 ETF on SET50 Index Spot-Futures Arbitrage Efficiency[1]

*Anant Chiarawongse**

Abstract

In 2007, Stock Exchange of Thailand lists ThaiDEX SET50 Exchange Traded Fund (TDEX), its first equity Exchange Traded Fund (ETF) that is designed to track SET50 index. Futures contracts with SET50 index as the underlying asset are in existent since 2006. This paper studies an impact such a listing of TDEX might have on index arbitrage activity. In a market without friction, mispricing between futures and the underlying index should be eliminated instantaneously by arbitrageurs. As a result, futures prices and index level will have a certain relationship according to the cost-of-carry model. However, in a real market situation, what seems like mispricing might persist due to limits to arbitrage. Since the introduction of TDEX can potentially relaxes some of those limits, we expect the mispricing to be less persistent after the introduction of TDEX.

To quantitatively measure the responsiveness that prices adjust to mispricing, we apply an error correction model to daily data of SET50 index futures and SET50 index. The coefficient of the error term in the model, which is called speed of adjustment coefficient, is a natural measure of speed that prices adjust to reach equilibrium. If the availability of TDEX helps facilitating arbitrage transactions, then the speed of adjustment coefficient is expected to increase after the listing of TDEX.

Keywords: Arbitrage, ETF, Cointegration, Error Correction Model

---

[1] The author wish to thank research support from the Ratchadaphiseksomphot Endowment Fund of Chulalongkorn University (RES560530092-HS). Email address: anant@cbs.chula.ac.th

* Lecturer, Department of Banking and Finance, Faculty of Commerce and Accountancy, Chulalongkorn University.

ผลกระทบของการเปิดให้ซื้อขายไทยเด็กซ์ เซ็ท 50 อีทีเอฟ ต่อประสิทธิภาพการทำอาบิทราจระหว่างดัชนีเซ็ท 50 และสัญญาซื้อขายล่วงหน้าที่มีดัชนีเซ็ท 50 เป็นสินค้าอ้างอิง[1]

*อนันต์ เจียรวงศ์**

บทคัดย่อ

ในปี พ.ศ. 2550 ตลาดหลักทรัพย์แห่งประเทศไทยได้เปิดให้ซื้อขาย ThaiDEX SET50 ซึ่งเป็นกองทุนรวมที่จดทะเบียนซื้อขายในตลาดหลักทรัพย์ (อีทีเอฟ) ที่ออกแบบให้มีผลตอบแทนตามดัชนี SET50 สัญญาฟิวเจอร์ที่อ้างอิงดัชนี SET50 เช่นเดียวกัน ได้เปิดให้ซื้อขายตั้งแต่ปี พ.ศ. 2549 บทความนี้ศึกษาว่าการเปิดให้ซื้อขาย TDEX มีผลกระทบต่อการทำอาบิทราจบนดัชนีหุ้นอย่างไร ในตลาดที่ไม่มีความฝืด ส่วนต่างราคาที่ผิดปรกติระหว่างฟิวเจอร์และดัชนีอ้างอิงควรจะถูกทำกำไรทันทีโดยผู้ทำอาบิทราจ ดังนั้นราคาฟิวเจอร์และระดับดัชนีควรจะมีความสัมพันธ์ที่เป็นไปตามตัวแบบต้นทุนการถือครอง แต่ในสภาพตลาดจริง ส่วนต่างราคาที่ผิดปรกติอาจจะคงอยู่ได้เป็นเวลานาน อันเป็นผลจากข้อจำกัดในการทำอาบิทราจ เนื่องจากการเปิดให้ซื้อขาย TDEX อาจช่วยลดข้อจำกัดในการทำอาบิทราจให้น้อยลง ผู้เขียนคาดหมายว่าส่วนต่างราคาที่ผิดปรกติจะคงอยู่ได้สั้นลงหลังจากการเปิดให้ซื้อขาย TDEX

ผู้เขียนใช้ตัวแบบตอบสนองความคลาดเคลื่อนกับข้อมูลราคารายวันของสัญญาฟิวเจอร์บนดัชนี SET50 และค่าของดัชนี SET50 ในการวัดการตอบสนองของราคาต่อส่วนต่างราคาที่ผิดปรกติ ค่าสัมประสิทธิ์ของพจน์ค่าความคลาดเคลื่อนซึ่งมีชื่อเรียกว่า สัมประสิทธิ์ความเร็วในการปรับตัว เป็นตัววัดความเร็วที่ราคาปรับตัวเข้าสู่ค่าที่ควรจะเป็น ดังนั้น ถ้าการมี TDEX ช่วยให้ทำอาบิทราจได้ง่ายขึ้น ผู้เขียนคาดหมายว่า สัมประสิทธิ์ความเร็วในการปรับตัวควรจะมีค่าเพิ่มขึ้นหลังจากการเปิดให้ซื้อขาย TDEX

คำสำคัญ: อาบิทราจ อีทีเอฟ โคอินทิเกรชั่น ตัวแบบตอบสนองความคลาดเคลื่อน

[1] ผู้เขียนขอขอบคุณทุนสนับสนุนการวิจัยจากกองทุนรัชดาภิเษกสมโภชจุฬาลงกรณ์มหาวิทยาลัย (RES560530092-HS)

* อาจารย์ประจำภาควิชาการธนาคารและการเงิน คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

1. Introduction

On April 28, 2006, the first ever stock index futures contract with SET50 index as the underlying starts trading on Stock Exchange of Thailand (SET). Apart from hedging and speculation, index futures contracts may be used in index arbitrage transactions. A pure index arbitrage involves buying (selling) the spot index via a basket trading and simultaneously selling (buying) futures contracts. However, index arbitrage on SET is particularly difficult in comparison to a mature market such as U.S. market because of higher margin requirement, high transaction cost, short-sale restriction, and illiquidity of index constituents that are small stocks. All these reasons combined renders index arbitrage on SET very costly, at least for retail investors. As a result, SET50 futures price and SET50 index level might not exhibit strong relationship as predicted by no arbitrage principle.

In 2007, SET lists its first exchange traded fund, ThaiDEX SET50 Exchange Traded Fund (TDEX). It is interesting to inquire what impacts the introduction of an exchange traded fund (ETF) such as TDEX might have on related products, i.e., SET50 index and SET50 futures. The introduction of TDEX not only gives institutional investors an additional instrument for index arbitrage operation, but also opens up the opportunity for retail investors to participate in index arbitrage activities as well. As a result, the deviation of futures price from its theoretical value should become more short-lived, i.e., the futures prices should adjust to equilibrium levels at a faster speed after the introduction of the TDEX.

The organization of this paper is as follow, section 2 provides a literature review of ETF related research. The data and methodology are described in section 3 and 4, respectively. Section 5 reports the results. The last section concludes.

2. Literature review

Gorton and Pennacchi (1993) provide a theoretical rationale for the packaging of seemingly redundant composite securities, i.e., securities whose value are aggregates of other assets' values. These prepackaged composite securities have lower variance of return in comparison to a basket of securities constructed by each trader. ETF is one example of such composite securities. An ETF is potentially useful to arbitrageurs, since it can simplify and speed up the execution of the cash leg of arbitrage transactions. Prior to ETF availability, to capture arbitrage opportunity, an arbitrageur has to trade the cash leg by buying or selling a basket of all stocks that comprise the index with the same weights as the index composition, and simultaneously takes a position in the futures contract in the opposite direction. Even with the invent of program trading, buying and selling security baskets is still difficult, because stocks with small weights are likely to result in an odd lot. These stocks with small weight are also likely to have the lowest liquidity. To mitigate the problem, an arbitrageur may opt for a basket with fewer securities at the cost of imprecisely replicating the index, thus introducing tracking error (Sofianos, 1993).

A single order in ETF achieves the same effect that requires orders in all stocks that constituting the index. Thus, the cash leg can be done faster and cheaper with ETF as opposed to a basket of securities. This faster execution will help driving futures mispricing back to the equilibrium value at faster speed; thus, arbitrage opportunities at any time are expected to be less persistent in the market.

Some papers examine pricing efficiency of ETF itself, e.g., Ackert and Tian (2000), who study whether the price of SPDRs ETF is close to fundamental price. Some papers investigate the impact of introducing ETF on underlying stocks, e.g., Hegde and McDermott (2004) who study the impact of introducing DIAMONDS ETF and QQQ ETF on the liquidity of underlying stocks, or Richie and Madura (2007) who study the impact of introducing QQQ on the liquidity and risk of underlying stocks.

Another group of papers study the impact of ETF introduction on the spot-futures pricing relationship. For example, Kurov and Lasser (2002) using tick-by-tick Nasdaq-100 futures transactions data conclude that the introduction of QQQ ETF help improving futures pricing efficiency. Switzer, Varson, and Zghidi (2000) using intraday data find that after the introduction of SPDR ETF, deviation of S&P 500 futures prices from theoretical price is also reduced. Related, Richie, Daigler, and Gleason (2008), using minute-by-minute prices of the S&P 500 futures, examine whether the SPDR ETF instead of baskets of stocks that replicate S&P 500 cash index can circumvent limits to arbitrage inherent in the latter. The results indicate that using SPDR in place of the cash index mitigate some limits to arbitrage, namely, trading costs, staleness of the underlying cash index, and volatility associated with the staleness. However, SPDR does have a limit of its own, namely, limited volume size.

In this paper, we investigate the impact that the introduction of ETF might have on the pricing efficiency of futures contract relative to the cash index. Since it is costly to do arbitrage in Thailand, we expect that if ETF can help circumventing limits to arbitrage, we should find more dramatic improvement on the arbitrage efficiency than other markets that are less costly.

The act of arbitrageurs will keep the diffierence between spot and futures prices in accord with cost-of-carry model. If arbitrage is at work, at a minimum we would expect the futures prices and cash index to move together in tandem. In other words, they are cointegrated, provided that the time series of both futures prices and cash index level are integrated of the same order. If arbitrage can't be done at all before the introduction of ETF, then the futures prices may not be cointegrated with cash index level. If the introduction of ETF makes arbitrage possible, we might observe cointegration between the two series. If we do not observe cointegration before the introduction of TDEX, but do observe cointegration after, then we may conclude that the introduction of ETF potentially improves the pricing efficiency of futures. However, if the prices are cointegrated both before and after the introduction of TDEX, then the impact of TDEX introduction is not identifiable by this before-and-after comparison, and further steps are needed.

When futures prices and spot index are cointegrated, there exists a representation called an error correction model. An error correction model contains an error term which can be interpreted as a deviation of price levels from the long run equilibrium level. The coefficient of error term naturally measures the speed of adjustment that prices move toward long-run equilibrium level. Error correction models have been used to determine lead-lag relationship between stock index cash market and futures market. For example, Shyy, Vijayraghavan, and Scott-Quinn (1996) find that CAC 40 futures price leads the CAC40 cash index, Judge and Reancharoen (2014) find that SET50 cash index leads the SET50 futures price, while Wahab and Lashgari (1993) studying S&P 500 and FTSE 100, Pizzi, Economopoulos, and O'Neill (1998) studying

S&P500, and Kavussanos, Visvikis, and Alexakis (2008) studying FTSE/ATHEX-20 all find bidirectional relationship between cash index and futures price.

The focus of this paper, however, is on the interpretation of the coefficient of error correction term. Our motivation is that it is far from clear how to measure arbitrage efficiency in the long run. Papers that rely on intraday data are by nature restricted to short-run. We propose to use the error correction coefficient as a potential measure of arbitrage efficiency in the long run. Larger coefficient means markets can impound information faster. As such, we expect the coefficient in the period before the introduction of TDEX to be smaller in magnitude than that in the period after the introduction of TDEX. We deem this novel interpretation of the error correction coefficient as a contribution of this paper.

Note that we focus on the pricing efficiency between futures contract and cash index because we employ pre-and-post analysis to assess the impact of ETF introduction, where we measure the impact by the coefficient of error correction term in an error correction model between two instruments that form an arbitrage pair. Employing pre-and-post analysis does impose a constraint on the data availability. We need time series of both instruments to exist before the introduction of ETF. This rules out ETF itself. SET50 index option is also ruled out because it is not yet available at the time of ETF introduction.

3. Data

3.1. Product details

SET50 index is a market capitalization weighted index, its value is based on 50 stocks with highest market capitalization and high level of liquidity. The exact criteria for choosing the 50 stocks are as follow. First, 200 stocks are chosen, these are stocks that have been listed for at least 6 months and have highest average daily market capitalization over the past 3 months. The list is further narrowed down by the requirement that stocks must have at least 20% free float and must have steady high level of liquidity. From the final list, the top 50 stocks are chosen to be in SET50. The list is reviewed every 6 months. Stocks that fail some of the aforementioned criteria are replaced by the next ones on the list. As of May 2014, the market capitalization of SET50 is 72.60% of the SET index, which is taken as representing market capitalization as a whole.[2]

SET50 futures contract is the first futures contract listed on SET. It has SET50 index as the underlying asset. The first trade date is April 28, 2006. At any time, there are six contract maturities, consisting of three nearest consecutive months, and next three quarterly months. The quarterly months are March, June, September, and December. The contract specification is as follow: contract multiplier is 1,000 baht per an index point, the minimum price fluctuation is 0.1 index point or 100 Baht. On May 6, 2014, SET50 futures contract specification is revised to become mini SET50 futures: a former SET50 futures contract is now equivalent to 5 mini SET50 futures contracts. To avoid complication from contract splitting, we restrict our data period of both SET50 cash index and SET50 futures to the period between April 28, 2006 and May 2, 2014.[3]

An exchange traded fund is an open-ended fund that is listed on an exchange and thus can be traded like a stock. The first exchange traded fund listed on SET is ThaiDEX SET50 exchange traded fund (TDEX), its first trading day is September 6, 2007. TDEX is a passive fund with the objective to replicate the return of SET50 Index. Later SET lists many more ETFs, but TDEX is still the largest and the most active ETFs on SET.[4]

Shares of ETF can be created or redeemed by Participating Dealers, which are typically large institutional investors such as a broker-dealers. Participating Dealers can send purchase/sale order directly to the company that manages ETF. In the case of TDEX, the minimum purchase order is 1 creation unit or 2,000,000 investment units. When an ETF is trading at a premium to its net asset value, participating dealers will have an incentive to capture the spread between the fund’s NAV and its market price. They can achieve this by buying the underlying components that constitute a creation basket and exchange it with the fund for new shares of the ETF. If ETF is trading at a discount to its net asset value, the reverse incentive applies. Thus, the price of ETF can be expected to track the underlying index well.

[2] http://www.set.or.th/th/market/market_statistics.html

[3] May 3 and 4 are Saturday and Sunday, while May 5 is a public holiday.

[4] ETF Monthly Report, May 2014.

3.2. Trading cost comparison

In this section, we compare the trading costs of a basket of stocks that replicate the cash index to the trading cost of TDEX ETF. A direct trading cost is commission fee. The commission fee for trading in ordinary stock is 0.25% for phone order or 0.15% for internet trading, while the commission fee for TDEX is only 0.1%. Another component of trading cost is bid-ask spread. For trading in ordinary stocks, there is a minimum spread requirement that depends on the price of stock as shown in table 1.

Table 1 Minimum spread at each price range

Market price level	Spread
less than 2 Baht	.01 Baht
from 2 baht but less than 5 Baht	.02 Baht
from 5 baht but less than 10 Baht	.05 Baht
from 10 baht but less than 25 Baht	.10 Baht
from 25 baht but less than 100 Baht	.25 Baht
from 100 baht but less than 200 Baht	.50 Baht
from 200 baht but less than 400 Baht	1.00 Baht
from 400 baht and more	2.00 Baht

From the table, the spread is at least 0.25% (1/400) when the price is between 200 and 400 Baht, and could be as high as 1% (.02/2) when the price is between 2 and 5 Baht In contrast, the minimum spread for TDEX is fixed at 0.01 Baht. Therefore, as long as TDEX prices stay above 4 Baht, the commission fee will be lower than 0.25%. At the inception price of 5.88 Baht per share for TDEX, the minimum spread is only 0.17%. However, during the period from October 6, 2008 to June 18, 2009, TDEX prices do get below 4 Baht and reach the lowest level at 2.74 Baht on October 27, 2008.

Another implicit cost of trading is short-sale restriction. The stocks that are eligible for short-selling are SET100 constituents and ETFs. Stocks must be available for borrowing and lending as well. In addition, short selling transactions are subject to uptick rule, i.e., the order price must be at a higher price than the last trade price.

Thus short selling ETF is much less restrictive than short selling a basket of 50 stocks. The combination of all costs indicates the potential of TDEX to be used in the cash legs of arbitrage transactions in Thailand.

3.3. Data

We use daily data on SET50 cash index and SET50 futures covering the period from the introduction of SET50 index futures (April 28, 2006) to the last day before the futures contracts specification is changed to mini future contracts. We don't use the TDEX data per se except that we use the date that TDEX is first traded, September 6, 2007, as a break point to divide the SET50 cash index and SET50 futures data into two periods: Pre-TDEX from April 28, 2006 to September 5, 2007, and Post-TDEX from September 6, 2007 to May 2, 2014. The data are extracted from DataStream. The futures prices are from continuous contract series, constructed from concatenating prices of nearest month contracts.

Table 2 Summary statistics of SET50 index, SET50 futures, and ThaiDEX ETF. The data frequency is daily. TDEX data is available from September 6, 2007 onward

	4/28/2006-5/2/2014		Pre-TDEX 4/28/2006-9/5/2007		Post-TDEX 9/6/2007-5/2/2014		
	Index	Futures	Index	Futures	Index	Futures	TDEX
No. Obs.	1958	1958	332	332	1626	1626	1626
Minimum	261.30	258.30	428.73	427.60	261.30	258.30	2.74
Maximum	1092.27	1097.90	638.56	632.80	1092.27	1097.90	11.04
Mean	651.39	648.75	504.99	504.82	681.28	678.14	6.87
Median	610.35	610.95	492.79	493.35	684.91	681.10	6.86
Stdev	200.19	200.65	42.26	41.54	206.47	207.46	2.08
Skewness	0.25	0.23	1.12	0.99	-0.08	-0.09	-0.05
Kurtosis	-0.88	-0.86	0.84	0.54	-0.90	-0.88	-0.94

Table 2 contains summary statistics of the data. The whole sample contains 1958 observations. Pre-TDEX period has only cash index level and futures price, while Post-TDEX period has an additional column containing the data on TDEX. During the whole period, the SET index reaches the minimum at 261.30, while the maximum at 1092.27 is around 4 times of the minimum. The negative kurtosis indicates that the index has a distribution with thinner tails than a normal distribution. Given the negative kurtosis, a positive skewness with small magnitude indicates that the distribution is right skew to a small degree. The summary statistics of futures prices are similar. However, by comparison the data of Pre-TDEX period and Post-TDEX period are starkly different. Pre-TDEX period data exhibit positive skewness and kurtosis, while post-TDEX data exhibit negative skewness and kurtosis. Taken together, the data exhibit considerable variability.

4. Methodology

Theoretical value of forward prices are related to the spot index level through a cost-of-carry model. Relation between futures price and spot index level is more complicate due to the marking-to-market of futures contract at the end of day. Brenner and Kroner (1995) relying on a work of Amin and Jarrow (1991), use a no-arbitrage, cost-of-carry asset pricing model to derive the following relationship between spot index and futures price:

$$ln\, F_{t,T} \;=\; ln\, S_t + \; ln\, D_{t,T} \;+\; Q_{t,T}$$

where $F_{t,T}$ is the price at time t of a futures contract that expires at time T, $D_{t,T} = r_{t,T} - V_{t,T}$ is the cost-of-carry term, which is the difference between spot interest rate $r_{t,T}$ of a pure discount bond maturing at time T, and continuous dividend yield $V_{t,T}$ over the period from t to T, and $Q_{t,T}$ is an adjustment term for the marking-to-market which depends on the volatility of interest rate and volatility of spot index. In the follow, we define $s_t = \ln S_t$ and $f_t = \ln F_{t,T}$

The cost-of-carry model determines fair futures price relative to spot index. Behind the cost-of-carry model is a no-arbitrage assumption. Whether s_t and f_t

behave according to the cost-of-carry model or not depending on arbitrage efficiency. If arbitrageurs can capture arbitrage opportunity then s_t and f_t will have to move together, even if s_t and f_t are nonstationary. A formal concept that describes long-run equilibrium relation between non-stationary time series is cointegration. Thus, testing whether s_t and f_t are cointegrated is akin to testing how well arbitrage works. Markets without friction will have the futures prices cointegrated with cash index level.

4.1. Unit roots test

A cointegration relation is a relation between nonstationary time series. Therefore, the first step is to determine whether f_t and s_t are nonstationary, in other words, containing unit roots. Nonstationary time series can be made stationary by differencing. A nonstationary time series that becomes stationary after differencing *d* times is said to be integrated of order *d,* written as *I(d).* To test the order of integration, we follow the sequential steps in applying Augmented Dickey Fuller (ADF) statistics and *t*-tests as outlined by Perron (1988).

Let y_t be the time series data of the variable of interest. Since the data generating process is not known, we start with an encompassing specification that contains both a drift term and a time trend term:

$$\Delta y_t = \mu + \gamma t + (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t \qquad (1)$$

The term $\Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j}$ is to ensure that there is no serial correlation in the error, the number of lags k is chosen by Akaike Information Criteria (AIC). The drift term μ is included for the possibility that y_0 might be different from 0, and γt represents a deterministic time trend. The steps to test for unit roots are as follows:

1. Test $H_0: \rho - 1 = 0$ against $H_1: \rho - 1 < 0$. The appropriate test statistic is not the usual *t* statistic but is a test statistic denoted by τ_τ, whose critical values are provided in table 8.5.2 of Fuller (1976).
   (a) If H_0 is not rejected, go to step 2.

(b) If H_0 is rejected, then the process is stationary.

2. Given that the process may contain a unit root, test for the presence of a trend: $H_0: \rho - 1 = 0,\ \gamma = 0$. The appropriate test statistic is Φ_3, critical values of which are provided in table VI of Dickey and Fuller (1981).

(a) If H_0 is not rejected, go to step 3.

(b) If H_0 is rejected, then proceed to test whether $\rho - 1 = 0$ given that γ is not zero. The test statistic is t statistics. If $\rho - 1 = 0$ is rejected, then the process is stationary, else we conclude that there is a unit root.

3. Drop the time trend by setting γ = 0, and estimate the regression equation:

$$\Delta y_t = \mu + (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t \qquad (2)$$

Then test $H_0: \rho - 1 = 0$ against $H_1: \rho - 1 < 0$. The test statistics is τ_μ, critical values are provided in table 8.5.2 of Fuller (1976).

(a) If H_0 is not rejected, go to step 4.

(b) If H_0 is rejected, then the process is stationary.

4. Test for the presence of a drift: $H_0: \rho - 1 = 0,\ \mu = 0$. The test statistic is Φ_1, critical values are provided in Dickey and Fuller (1981), table IV.

(a) If H_0 is not rejected, go to step 5.

(b) If H_0 is rejected, then proceed to test whether $\rho - 1 = 0$ given that μ is not zero. The test statistic is t statistics. If $\rho - 1 = 0$ is rejected, then the process is stationary, else we conclude that there is a unit root.

5. Set $\mu = 0$, and estimate the regression equation:

$$\Delta y_t = (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t \qquad (3)$$

Then test $H_0: \rho - 1 = 0$ against $H_1: \rho - 1 < 0$. The test statistics is τ, critical values are provided in table 8.5.2 of Fuller (1976). If H_0 is rejected, then the process is stationary, else we conclude that there is a unit root.

If the test indicates the presence of unit roots, we can only conclude that the process is integrated of order 1 or higher. To determine whether the process is *I*(1), we replace y_t with its first difference, Δy_t and test for unit roots. If the result indicates that Δy_t is stationary, then we conclude that y_t is *I* (1).

4.2. Cointegration

Provided that f_t and s_t are nonstationary, we can proceed to test for cointegration. We first briefly explain the concept of cointegration. Consider two univariate time series x_t and y_t. If both are *I*(d), then x_t and y_t are said to be cointegrated of order *d, b* if there is a vector α such that $z_t \equiv \alpha' \cdot \begin{bmatrix} x_t \\ y_t \end{bmatrix}$. Cointegration represents a long-run equilibrium relation between x_t and y_t

Engle and Granger (1987) propose a two-step procedure to estimate a cointegrating vector. In our context, suppose that both f_t and s_t are *I* (1). The first step involves running a linear regression between f_t and s_t

$$f_t = a + bs_t + z_t \qquad (4)$$

If the two times series are indeed cointegrating, then the z_t must be stationary. Therefore, the second step tests whether the time series of the residuals $\hat{z}_t$ contains a unit root. If the residuals contain no unit roots, we can conclude that f_t and s_t are cointegrated. Note that the residual $\hat{z}_t$ has zero mean because an intercept term is included in equation 4. Thus, to test for unit root, we run the regression equation without the drift term:

$$\Delta\hat{z}_t = (\rho - 1)\hat{z}_{t-1} + \Sigma_{j=1}^{k}\phi_j \Delta\hat{z}_{t-j} + \nu_t \qquad (5)$$

A time trend is not included either because, according to Hansen (1992), including a time trend results in a loss of power. The lags of $\hat{z}_t$ are included to purge the serial correlation in the error, the appropriate number of lags k is determined by AIC. The critical values are computed via the response surface given in MacKinnon (2010) which is an updated version of MacKinnon (1991). Using MacKinnon critical values is more accurate because they are adjusted for the number of time series in the cointegrating relation and for the number of observations as well.

We will test whether f_t and s_t are cointegrated in both the Pre-TDEX and Post-TDEX periods

4.3. Error correction model

If f_t and s_t are cointegrated in both periods, then we cannot conclude whether the introduction of TDEX has any impact on pricing efficiency of futures contract relative to cash index. However, we may be able to detect the impact through short-run dynamics which explicitly describe how variables of interest response to deviation in the short run. In fact, from Engle-Granger two-step approach, we obtain $\hat{z}_t = f_t - \hat{a} - \hat{b}s_t$ which can be interpreted as a deviation from the long run equilibrium. According to Granger's Representation Theorem (Engle & Granger, 1987), a cointegrated system has an equivalent representation called error correction model. The bivariate case that we consider has the following error correction representation:

$$\Delta f_t = \alpha_f + \gamma_f \hat{z}_{t-1} + \Sigma_{i=1}^{K} \beta_{f,i} \Delta f_{t-i} + \Sigma_{i=1}^{L} \beta_{s,i} \Delta s_{t-i} + \varepsilon_{f,t} \qquad (6)$$

$$\Delta s_t = \alpha_s + \gamma_s \hat{z}_{t-1} + \Sigma_{i=1}^{M} \theta_{s,i} \Delta s_{t-i} + \Sigma_{i=1}^{N} \theta_{f,i} \Delta f_{t-i} + \varepsilon_{s,t} \qquad (7)$$

The coefficient γ_f measures the sensitivity of f_t to adjust in response to the deviation from long-run equilibrium $\hat{z}_t$. γ_s has a similar interpretation in terms of s_t. γ_f and γ_s measure speeds of adjustment to long-run equilibrium level of futures price and cash index, respectively. Error correction representation can be estimated by a linear regression because all the variables are stationary.

Since TDEX can potentially substitute a basket of securities in index arbitrage transactions and help facilitating arbitrage transaction, we expect the arbitrage opportunity to be less persistent. This would reflect in a speed of adjustment coefficients with larger magnitude. Thus, to detect the impact of TDEX introduction on arbitrage efficiency, we estimate the error correction representation for Pre-TDEX and Post-TDEX periods separately and compare the speed of adjustment coefficients. We expect the magnitude of speed of adjustment coefficients to be larger after TDEX introduction.

5. Results

5.1. Unit roots test

We show the results of testing for the presence of unit roots for $s_t = \ln S_t$ in the Pre-TDEX period in details. The tests on other time series and the Post-TDEX period are similar.

Let $y_t = s_t$ during the Pre-TDEX period. We follow the steps outlined in Section 4.1. We first estimate the regression which include both drift and time trend as in equation 1. Using AIC, the value of lag k is chosen to be 1. The regression results and corresponding test statistics are shown in table 3.

Table 3 Unit root test on time series $s_t = \ln S_t$ for the Pre-TDEX period. This table is for the case where there may be a time-trend and a drift, corresponding to equation 1:

$$y_t = \mu + \gamma t + (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t$$

(a) Regression results. The lag k, determined by AIC, is 1.

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.2260	0.0870	2.60	0.0098
y_{t-1}	-0.0372	0.0142	-2.63	0.0090
t	0.0000	0.0000	2.83	0.0049
Δy_{t-1}	-0.1614	0.0546	-2.96	0.0033

(b) Test statistics and critical values. τ_τ is a test statistic for $H_0: \rho - 1 = 0,$ while Φ_3 is a test statistic for joint hypothesis $H_0: \rho - 1 = 0, \gamma = 0$

	Statistics	1pct	5pct	10pct
τ_τ	-2.63	-3.98	-3.42	-3.13
Φ_3	4.77	8.34	6.30	5.36

In step 1, we test $H_0: \rho - 1 = 0$. From table 3 panel (b), the value of test statistics τ_τ is -2.63, which is smaller in magnitude than critical values at every significance level. Thus, we cannot reject $H_0: \rho - 1 = 0$. We continue to step 2 where we test $H_0: \rho - 1 = 0, \gamma = 0$ The value of test statistic Φ_3 is 4.77, which is smaller than critical values at every level. Thus, we cannot reject the null hypothesis that there is a unit root without time trend. Dropping the time trend, we estimate equation 2. The regression results and corresponding test statistics are shown in table 4.

Table 4 Unit root test on time series $s_t = \ln S_t$, for the Pre-TDEX period. This table is for the case where there is no time-trend, but there is possibly a drift. The corresponding equation is equation 2: $\Delta y_t = \mu + (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t$

(a) Regression results. The lag k, determined by AIC, is 1.

	Estimate	**Std. Error**	***t*-value**	**Pr(>|t|)**
Intercept	0.0893	0.0732	1.22	0.2231
y_{t-1}	-0.0143	0.0118	-1.22	0.2244
Δy_{t-1}	-0.1612	0.0551	-2.92	0.0037

(b) Test statistics and critical values. τ_μ is a test statistic for $H_0: \rho - 1 = 0,$ while Φ_1 is a test statistic for joint hypothesis $H_0: \rho - 1 = 0, \mu = 0$.

	statistic	1pct	5pct	10pct
τ_μ	-1.22	-3.44	-2.87	-2.57
Φ_1	0.78	6.47	4.61	3.79

In step 3, we test $H_0: \rho - 1 = 0,$ where the appropriate test statistic is τ_μ Since the value of τ_μ is -1.22, which is smaller in magnitude than critical value at every significance level, we cannot reject the possibility of unit root. Then in step 4, we test the presence of unit root when the process has no drift $H_0: \rho - 1 = 0, \mu = 0$. The value of test statistic Φ_1 is 0.78, which is much smaller than critical value at every significance level. So we set $\mu = 0,$ and estimate equation 3. The regression results and corresponding test statistics are shown in table 5. Since the value of test statistic τ is only 0.26 which is smaller in magnitude than critical value at every significance level, we conclude that s_t contains a unit root.

Table 5 Unit root test on time series s_t =ln S_t, for the Pre-TDEX period. This table is for the case where there may be a time-trend and a drift, corresponding to equation 3: $\Delta y_t = (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t$

(a) Regression results. The lag $k,$ determined by AIC, is 1.

	Estimate	Std. Err	*t*-Value	Pr(>|*t*|)
y_{t-1}	0.0000	0.0002	0.26	0.7969
Δy_{t-1}	-0.1696	0.0548	-3.10	0.0021

(b) Test statistics and critical values. τ_μ is a test statistic for $H_0: \rho - 1 = 0$.

	statistic	1pct	5pct	10pct
τ	0.26	-2.58	-1.95	-1.62

It is possible that s_t is integrated of higher order than *I*(1). To determine whether the series is indeed *I*(1), we test whether differencing the series once makes the series stationary. So we repeat the unit root test but replace y_t with Δy_t. The results are shown in table 6. The value of τ_τ statistic is -13.36, which is overwhelmingly larger in magnitude than critical value at every significance level. Thus, we conclude that y_t is stationary. In other words, s_t is indeed *I*(1).

Table 6 Test to find the order of integration for the series s $s_t = \ln S_t$, Pre-TDEX period. The regression equation is equation 1 with $y_t = \ln S_t - \ln S_{t-1}$:
$$\Delta y_t = \mu + \gamma t + (\rho - 1)y_{t-1} + \Sigma_{j=1}^{k}\gamma_j \Delta y_{t-j} + u_t$$

(a) Regression results. The lag k, determined by AIC, is 1.

	Estimate	Std. Error	*t*-Value	Pr(>|*t*|)
Intercept	-0.0024	0.0019	-1.22	0.2223
y_{t-1}	-1.1436	0.0856	-13.36	0.0000
t	0.0000	0.0000	1.55	0.1226
$\Delta y_{t\ 1}$	-0.0276	0.0559	-0.49	0.6221

(b) Test statistics and critical values. τ_τ is a test statistic for $H_0: \rho - 1 = 0$, while is Φ_3 a test statistic for joint hypothesis $H_0: \rho - 1 = 0, \gamma = 0$.

	statistic	1pct	5pct	10pct
τ_τ	-13.36	-3.98	-3.42	-3.13
Φ_3	89.23	8.34	6.30	5.36

We repeat unit root test on the time series of s_t in the Post-TDEX period, and on f_t in both Pre-TDEX and Post-TDEX periods. The results indicate that all series are *I*(1). Thus, we can proceed to test whether f_t and s_t are cointegrated.

5.2. Cointegration

To test whether f_t and s_t are cointegrated in each of the Pre-TDEX and Post-TDEX periods, we run a linear regression in equation 4 to obtain the residuals $\hat{z}_t$. Then we test for the unit root by running the regression in equation 5. The results of cointegration test for the Pre-TDEX period are shown in table 7.

MacKinnon's critical values in panel (c) of table 7 are calculated from the following formula:

$$\hat{C}(p,T) = \hat{\beta}_\infty + \hat{\beta}_1 T^{-1} + \hat{\beta}_2 T^{-2}$$

T is a sample size. In our case, the original sample size for Pre-TDEX period is 332. The sample size when running equation 5 is 331 due to a lag operation. The parameters depend on the number of time series to be tested in cointegrating relationship, which is 2 in our case, significance level p, and whether the tested time series has a trend term or not. Since equation 5 does not include a time trend term, our case corresponds to "no trend" case in MacKinnon (2010).

At $p = 1$%, the values of parameter looked up from table 2 of MacKinnon $\hat{\beta}_\infty = -3.89644, \hat{\beta}_1 = -10.9519, \hat{\beta}_2 = -22.527$ For *T*=331, the critical value at 1% significance level is -3.93. Other critical values are calculated based on corresponding Parameters $\hat{\beta}_\infty, \hat{\beta}_1, \hat{\beta}_2$. From table 7, it turns out that we can reject the null that the residuals contain a unit root only at 10% significance level but cannot reject at 5% and *a fortiori* at 1% significance level. In other words, futures price and cash index appears to be cointegrated only at 10% significance level.

Table 7 Test of cointegration between ln S_t and ln F_t for the Pre-TDEX period.

(a) Regression coefficients from cointegrating equation 4, Pre-TDEX period: $f_t = a + bs_t + z_t$

	Estimate	Std. Error	*t*-value	Pr(>|*t*|)
Intercept	0.1163	0.0421	2.77	0.0060
s_t	0.9813	0.0068	145.14	0.0000

(b) Regression coefficients from running residuals regression, equation 5, Pre-TDEX period: $\Delta\hat{z}_t = (\rho - 1)\hat{z}_{t-1} + \Sigma_{j=1}^{k}\phi_j\Delta\hat{z}_{t-j} + \nu_t$

	Estimate	Std. Err	t-Value	Pr(>|t|)
y_{t-1}	-0.1211	0.0386	-3.13	0.0019
Δy_{t-1}	-0.3797	0.0603	-6.29	0.0000
Δy_{t-2}	-0.2872	0.0595	-4.83	0.0000
Δy_{t-3}	-0.1628	0.0549	-2.97	0.0032

(c) Test statistics for $H_0: \rho - 1 = 0$ in residual regression, Pre-TDEX period. The critical values are calculated by the response surface in MacKinnon (2010).

	statistic	1pct	5pct	10pct
τ	-3.13	-3.93	-3.35	-3.06

The results for Post-TDEX period are shown in table 8. The magnitude of test statistic exceeds the critical value at every significance level. Thus, we can conclude decisively that futures price and cash index are cointegrated after the introduction of TDEX.

Given the results of cointegration tests for both periods, it would be more informative to compare the results from a different angle, i.e., error correction model. An error correction representation exists for time series that are cointegrated. For the Post-TDEX period, the time series of f_t and s_t are cointegrated at every significance level, but only at 10% significance level for the Pre-TDEX period. For the Pre-TDEX period, we must be content with 10% significance level in order to proceed to estimate a corresponding error correction representation.

Table 8 Test of cointegration between ln S_t and ln F_t for the Post-TDEX period.

(a) Regression coefficients from cointegrating equation 4, Post-TDEX period $f_t = a + bs_t + z_t$

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	-0.1006	0.0039	-25.94	0.0000
s_t	1.0146	0.0006	1695.02	0.0000

(b) Regression coefficients from running residuals regression, equation 5, Post-TDEX period: $\Delta\hat{z}_t = (\rho - 1)\hat{z}_{t-1} + \Sigma_{j=1}^{k}\phi_j\Delta\hat{z}_{t-j} + \nu_t$

	Estimate	Std. Err	t-Value	Pr(>|t|)
y_{t-1}	-0.1258	0.0181	-6.94	0.0000
Δy_{t-1}	-0.4369	0.0279	-15.68	0.0000
Δy_{t-2}	-0.2293	0.0288	-7.95	0.0000
Δy_{t-3}	-0.1916	0.0280	-6.83	0.0000
Δy_{t-4}	-0.0765	0.0248	-3.09	0.0020

(c) Test statistics for $H_0: \rho - 1 = 0$ in residual regression, Post-TDEX period. The critical values are calculated by the response surface in MacKinnon (2010).

	statistic	1pct	5pct	10pct
τ	-6.94	-3.90	-3.34	-3.05

5.3. Error correction representation

We estimate error correction representation as in equation 6 and 7 for both Pre-TDEX period and Post-TDEX period. Table 9 panel (a) and (b) contains the results of estimating equation 6 and 7, respectively for the Pre-TDEX period. The lags of f_t and s_t are chosen such that there is no serial correlation in the error term. The variable of interest is $\hat{z}_{t-1}$, whose coefficient measures the speed of adjustment. For the equation that has f_t as the dependent variable, the speed of adjustment coefficient is -0.0939, and the associated p-value is 0.3125. When the dependent variable is s_t, the speed of adjustment coefficient is 0.0951, and the associated p-value is 0.3341. Therefore, for the Pre-TDEX period, both coefficients are neither economically significant nor statistically significant.

Table 9 Estimating error correction model, Pre-TDEX period.

(a) Results of estimating equation 6 with Δf_t as the dependent variable:

$$\Delta f_t = \alpha_f + \gamma_f \hat{z}_{t-1} + \Sigma_{i=1}^{K} \beta_{f,i} \Delta f_{t-i} + \Sigma_{i=1}^{L} \beta_{s,i} \Delta s_{t-i} + \varepsilon_{f,t}$$

	Estimate	Std. Err	t-Value	Pr(>|t|)
Intercept	0.0003	0.0009	0.35	0.7259
$\hat{z}_{t-1}$	-0.0939	0.0928	-1.01	0.3125
Δf_{t-1}	0.2623	0.1412	1.86	0.0641
Δs_{t-1}	-0.3211	0.1306	-2.46	0.0145

(b) Results of estimating equation 7 with Δs_t as the dependent variable:

$$\Delta s_t = \alpha_s + \gamma_s \hat{z}_{t-1} + \Sigma_{i=1}^{M} \theta_{s,i} \Delta s_{t-i} + \Sigma_{i=1}^{N} \theta_{f,i} \Delta f_{t-i} + \varepsilon_{s,t}$$

	Estimate	Std. Err	t-Value	Pr(>|t|)
Intercept	0.0003	0.0009	0.35	0.7267
$\hat{z}_{t-1}$	0.0951	0.0983	0.97	0.3341
Δf_{t-1}	0.4774	0.1496	3.19	0.0016
Δs_{t-1}	-0.5648	0.1384	-4.08	0.0001

Table 10 Estimating error correction model, Post-TDEX period.

(a) Results of estimating equation 6 with Δf_t as the dependent variable:

$$\Delta f_t = \alpha_f + \gamma_f \hat{z}_{t-1} + \Sigma_{i=1}^{K} \beta_{f,i} \Delta f_{t-i} + \Sigma_{i=1}^{L} \beta_{s,i} \Delta s_{t-i} + \varepsilon_{f,t}$$

	Estimate	Std. Err	-Value	Pr(>||)
Intercept	0.0003	0.0005	0.64	0.5213
$\hat{z}_{t-1}$	-0.2252	0.0626	-3.60	0.0003
Δf_{t-1}	-0.3872	0.0963	-4.02	0.0001
Δs_{t-1}	0.4198	0.1082	3.88	0.0001
Δf_{t-2}	-0.1251	0.0890	-1.41	0.1599
Δs_{t-2}	0.1676	0.0999	1.68	0.0935

(b) Results of estimating equation 7 with Δs_t as the dependent variable:

$$\Delta s_t = \alpha_s + \gamma_s \hat{z}_{t-1} + \Sigma_{i=1}^{M} \theta_{s,i} \Delta s_{t-i} + \Sigma_{i=1}^{N} \theta_{f,i} \Delta f_{t-i} + \varepsilon_{s,t}$$

	Estimate	Std. Err	t-Value	Pr(>|t|)
Intercept	0.0003	0.0004	0.73	0.4637
$\hat{z}_{t-1}$	-0.0664	0.0546	-1.21	0.2247
Δf_{t-1}	-0.0219	0.0841	-0.26	0.7949
Δs_{t-1}	0.0617	0.0945	0.65	0.5136
Δf_{t-2}	-0.0053	0.0777	-0.07	0.9458
Δs_{t-2}	0.0503	0.0872	0.58	0.5639

The results for Post-TDEX period are shown in table 10. The lags of f_t and s_t that purge serial correlation from the error term are 2. The speed of adjustment coefficients when f_t is the dependent variable is -.2252 which is much larger than that of Pre-TDEX period. It is also statistically significant with p-value at 0.0003 only. We interpret this result as an evidence that future prices adjust to $\hat{z}_{t-1}$ more responsively after TDEX becomes available. In contrast, the speed of adjustment coefficient when s_t is the dependent variable is -0.0664 and is not statistically significant. So it appears the introduction of TDEX doesn't have conclusive impact on the cash index. The results seem encouraging that TDEX may help facilitating arbitrage transaction. However, one might suspect whether the improvement in arbitrage efficiency stems from other factors that help improve market over time, e.g., from better order execution technology. Moreover, the sample sizes are different between the Pre-TDEX period and Post-TDEX period. There is a reason to suspect that the results may depends on the sample size. We address these two concerns in the next subsection

5.4. Robustness check

There are 1626 observations in Post-TDEX period. We divide the dataset into 5 sub-periods, each sub-period contains 325 observations, approximately the same size as the data set in the Pre-TDEX period. Then we repeat the analysis on each subset. First, we test whether time series, restricted to each subset, are cointegrated. If they are cointegrated, then we estimate the error correction representation. We report the result of cointegration test in table 11. Here we report only the test statistics for testing the hypothesis whether the residuals time series contains unit root.

Table 11 Test statistic for $H_0: \rho - 1 = 0$ from running equation 5 on Post-TDEX sub-period data: $\Delta\hat{z}_t = (\rho - 1)\hat{z}_{t-1} + \Sigma_{j=1}^{k}\phi_j\Delta\hat{z}_{t-j} + \nu_t$. The critical values calculated from MacKinnon (2010) are the same for all sub-periods, because all sub-periods have the same sample size.

Sub-period	From	To	Test statistic	1pct	5pct	10pct
1	09/06/2007	12/30/2008	-5.33			
2	01/05/2009	05/06/2010	-4.97			
3	05/07/2010	09/07/2011	-2.48	-3.93	-3.36	-3.06
4	09/08/2011	01/02/2013	-3.11			
5	01/03/2013	04/30/2014	4.33			

It turns out that the time series are cointegrated at 1% significance level in sub-periods 1, 2 and 5, and at 10% significance level in sub-period 4. We cannot reject the hypothesis that the time series are not cointegrated in sub-period 3. So, we estimate the error correction model only for periods 1, 2, 4 and 5. The results of error correction model estimation are summarized in table 12.

The results for sub-periods are similar to the whole period's results in the following sense. In each sub-period, the speed of adjustment coefficient when f_t is the dependent variable has large magnitude in comparison to the speed of adjustment coefficient when s_t is the dependent variable. In fact, all the speed of adjustment coefficients when s_t is the dependent variable are all close to zero, their p-value are also large. In other words, none of them are statistically and economically significant. All the speed of adjustment coefficients when f_t is the dependent variable are negative.

The first sub-period's coefficient of -0.4284 has the largest magnitude, while the second sub-period's coefficient of -0.1348 has the smallest magnitude. Therefore, we regard these coefficients as economically significant. In terms of statistically significance, only sub-periods 1 and 4 have p-values close to 5% significance level, while coefficients of sub-periods 2 and 5 are not statistically significant. Thus the results are not as strong as using the whole Post-TDEX period data. Taken all the results together, we conclude that the introduction of TDEX helps increase arbitrage efficiency, though not uniformly over time, as revealed by the speed of adjustment coefficients of the futures prices that are economically significant in all sub-periods and statistically significant at 10% level in sub-periods 1 and 4.

Table 12 Speed of adjustment coefficient from error correction model, equation 6 and 7, for sub-periods 1, 2, 4 and 5.

Sub-period	Dependent Variable	Speed of Adjustment	Std. Err	t-Value	Pr(>|t|)
1	Δf_t	-0.4284	0.2105	-2.04	0.0427
	Δs_t	-0.0097	0.1808	-0.05	0.9571
2	Δf_t	-0.1348	0.1217	-1.11	0.2689
	Δs_t	0.0569	0.1028	0.55	0.5807
4	Δf_t	-0.3970	0.2076	-1.91	0.0568
	Δs_t	-0.0446	0.1910	-0.23	0.8157
5	Δf_t	-0.3188	0.2443	-1.30	0.1928
	Δs_t	0.0582	0.2220	0.26	0.7933

Without resorting to a full multivariate analysis, we can perform a quick test by allowing the speed of adjustment coefficient to depend on another factor, namely TDEX volume. This is in line with Richie, Daigler, and Gleason (2008) who find that the effectiveness of using of SPDR ETF to replace cash index is limited by volume size. Letting Vol represent TDEX volume, we modify ECM as follow:

$$\Delta f_t = \alpha_f + \left(\gamma_{f0} + \gamma_{f1} Vol\right)\hat{z}_{t-1} + \Sigma_{i=1}^{K}\beta_{f,i}\Delta f_{t-i} + \Sigma_{i=1}^{L}\beta_{s,i}\Delta s_{t-i} + \varepsilon_{f,t} \quad (8)$$

$$\Delta s_t = \alpha_s + (\gamma_{s0} + \gamma_{s1} Vol)\hat{z}_{t-1} + \Sigma_{i=1}^{M}\theta_{s,i}\Delta s_{t-i} + \Sigma_{i=1}^{N}\theta_{f,i}\Delta f_{t-i} + \varepsilon_{s,t} \quad (9)$$

We apply these specifications to Post-TDEX period. The results are reported in table 13.

Table 13 Estimating error correction model with interactive term, Post-TDEX period.

(a) Results of estimating equation 8 with Δf_t as the dependent variable:

$$\Delta f_t = \alpha_f + \left(\gamma_{f0} + \gamma_{f1} Vol\right)\hat{z}_{t-1} + \Sigma_{i=1}^{K}\beta_{f,i}\Delta f_{t-i} + \Sigma_{i=1}^{L}\beta_{s,i}\Delta s_{t-i} + \varepsilon_{f,t}$$

	Estimate	Std. Err	*t*-Value	Pr(>|*t*|)
Intercept	0.0003	0.0005	0.68	0.4989
$\hat{z}_{t-1}$	-0.3124	0.1609	-1.92	0.0524
$Vol \cdot \hat{z}_{t-1}$	0.0009	0.0016	0.59	0.5566
Δf_{t-1}	-0.3845	0.0964	-3.99	0.0001
Δs_{t-1}	0.4174	0.1083	3.86	0.0001
Δf_{t-2}	-0.1221	0.0892	-1.37	0.1710
Δs_{t-2}	0.1644	0.1000	1.64	0.1005

(b) Results of estimating equation 9 with Δs_t as the dependent variable:

$$\Delta s_t = \alpha_s + (\gamma_{s0} + \gamma_{s1} Vol)\hat{z}_{t-1} + \Sigma_{i=1}^{M}\theta_{s,i}\Delta s_{t-i} + \Sigma_{i=1}^{N}\theta_{f,i}\Delta f_{t-i} + \varepsilon_{s,t}$$

	Estimate	Std. Err	*t*-Value	Pr(>|*t*|)
Intercept	0.0003	0.0004	0.77	0.444
$\hat{z}_{t-1}$	-0.1372	0.1405	-0.98	0.329
$Vol \cdot \hat{z}_{t-1}$	0.0008	0.0014	0.55	0.584
Δf_{t-1}	-0.0197	0.0842	-0.23	0.815
Δs_{t-1}	0.0597	0.0946	0.63	0.528
Δf_{t-2}	-0.0028	0.0779	-0.04	0.971
Δs_{t-2}	0.0477	0.0874	0.55	0.585

We focus on the coefficients of the interactive term Vol $\cdot\, \hat{z}_{t-1}$. For equation 8, the coefficient is 0.0009 with p-value 0.5566, while for equation 9, the coefficient is 0.0008 with p-value 0.584. So it turns out that TDEX volume has very small effect on the speed of adjustment coefficient and is not statistically significant. The results are surprising as they indicate that the improvement in arbitrage efficiency does not follow directly from using TDEX to replace the cash index. However, the results are similar to those reported by Deville, Gresse, and Séverac (2014) who study CAC40 index. Therefore the source of efficiency after the introduction of TDEX for Stock Exchange of Thailand is yet to be determined.

6. Conclusions

In this paper we set out to investigate the impact that the listing of a new financial product such as TDEX might have on the related product, i.e., SET50 futures contract. Both products have SET50 as the reference asset. The premise is that a product such as TDEX has a number of desirable characteristics that make it a good substitute for the cash leg in index arbitrage transactions. Thus, it should help improve arbitrage efficiency as a consequence. We propose to measure the arbitrage efficiency by a speed of adjustment coefficient from an error correction model and expect the speed of adjustment coefficient to increase after the listing of TDEX. Dividing the data into Pre-TDEX period and Post-TDEX period, we estimate the speed of adjustment coefficients. There are two dependent variables, the change in log futures price and the change in log index level. It turns out that the change in log index level is not much responsive to the error term in both periods. The change in log futures level, however, responses much faster to the error term after the introduction of TDEX. So it appears that TDEX does help improving arbitrage efficiency.

We investigate further by dividing the Post-TDEX period into five sub-periods so that the number of observations in each sub-period is approximately the same as the Pre-TDEX period, and repeat the test of cointegration and estimation of error correction model for each sub-period. Judging from the speed of adjustment coefficients of futures prices, TDEX appears to help improving arbitrage efficiency in all sub-periods, but the coefficients are statistically significant only in sub-periods 1 and 4. We probe further by allowing the speed of adjustment coefficient to depend on the volume of TDEX. However, the results indicate that TDEX volume are not economically or statistically significant in determining the speed of adjustment coefficient. We suggest that determining sources of arbitrage efficiency after the introduction of TDEX would be an interesting research topic.

References

Ackert, L. F. & Tian, Y. S. (2000). Arbitrage and valuation in the market for Standard and Poor's depositary receipts. ***Financial Management***, *29*(3), 71–87.

Amin, K. I. & Jarrow, R. A. (1991). Pricing foreign currency options under stochastic interest rates. ***Journal of International Money and Finance***, *10*(3), 310–329.

Brenner, R. J. & Kroner, K. F. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. ***The Journal of Financial and Quantitative Analysis***, *30*(1), 23–42.

Deville, L., Gresse, C., & Séverac, B. (2014). Direct and indirect effects of index ETFs on spot-futures pricing and liquidity: evidence from the CAC 40 index. ***European Financial Management,*** *20*(2), 352-373.

Dickey, D. A. & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. ***Econometrica***, *49*(4), 1057–1072.

Engle, R. F. & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. ***Econometrica***, *55*(2), 251–276.

Fuller, W. A. (1976). ***Introduction to Statistical Time Series***. New York, NY: John Wiley & Sons, Inc.

Gorton, G. B. & Pennacchi, G. G. (1993). Security baskets and index-linked securities. ***The Journal of Business***, *66*(1), 1–27.

Hansen, B. E. (1992). Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends. ***Journal of Econometrics***, *53*(1), 87–121.

Hegde, S. P. & McDermott, J. B. (2004). The market liquidity of DIAMONDS, Q's, and their underlying stocks. ***Journal of Banking & Finance***, *28*(5), 1043–1067.

Judge, A. & Reancharoen, T. (2014). An empirical examination of the lead-lag relationship between spot and futures markets: Evidence from Thailand. ***Pacific-Basin Finance Journal***, *29*, 335-358.

Kavussanos, M. G., Visvikis, I. D., & Alexakis, P. D. (2008). The lead-lag relationship between cash and stock index futures in a new market. ***European Financial Management***, *14*(5), 1007-1025.

Kurov, A. A. & Lasser, D. J. (2002). The effect of the introduction of cubes on the NASDAQ-100 index spot-futures pricing relationship. ***The Journal of Futures Markets***, *22*(3), 197–218.

MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle and C. W. J. Granger (Eds.), ***Long-Run Economic Relationships: Readings in Cointegration***, (pp. 267-276). Oxford: Oxford University Press.

MacKinnon, J. G. (2010). ***Critical values for cointegration tests***. (Working Paper No. 1227), Retrieved from http://qed.econ.queensu.ca/working_papers/papers/qed_wp_1227.pdf

Perron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach. ***Journal of Economic Dynamics and Control***, *12*(2–3), 297 – 332.

Pizzi, M. A., Economopoulos, A. J., & O'Neill, H. M. (1998). An examination of the relationship between stock index cash and futures markets: a cointegration approach. ***The Journal of Futures Markets***, *18*(3), 297-305.

Richie, N., Daigler, R. T., & Gleason, K. C. (2008). The limits to stock index arbitrage: Examining S&P 500 futures and SPDRS. ***The Journal of Futures Markets***, *28*(12), 1182–1205.

Richie, N. & Madura, J. (2007). Impact of the QQQ on liquidity and risk of the underlying stocks. ***The Quarterly Review of Economics and Finance***, *47*(3), 411–421.

Shyy, G., Vijayraghavan, V., & Scott-Quinn, B. (1996). A further investigation of the lead-lag relationship between the cash market and stock index futures with the use of bid/ask quotes: the case of France. ***The Journal of Futures Markets***, 16(4), 405-420.

Sofianos, G. (1993). Index arbitrage profitability. ***Journal of Derivatives***, *1*(1), 6–20.

Switzer, L. N., Varson, P. L., & Zghidi, S. (2000). Standard and Poor's depository receipts and the performance of the S&P 500 index futures market. ***The Journal of Futures Markets***, *20*(8), 705–716.

Wahab, M., & Lashgari, M. (1993). Price dynamics and error correction in stock index and stock index futures markets: a cointegration approach. ***The Journal of Futures Markets***, 13(7), 711-742.