Distribution Agreement

Index Spot-Futures Arbitrage: Evidence from Examining S&P 500 Index and SPDR ETF

By

Jianhao Zhang

Caroline Fohlin
Adviser

Economics

Caroline Fohlin
Adviser

Le Chen
Committee Member

Elena Pesavento
Committee Member

2021

Index Spot-Futures Arbitrage: Evidence from Examining S&P 500 Index and SPDR ETF

By

Jianhao Zhang

Caroline Fohlin
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2021

Abstract

Index Spot-Futures Arbitrage: Evidence from Examining S&P 500 Index and SPDR ETF

By Jianhao Zhang

This paper aims to examine the mispricing or arbitrage opportunities between the spot-futures relationship of the S&P 500 index. We use minute-by-minute intraday E-mini S&P 500 Futures trading data to explore the frequency and magnitude of mispricing with respect to two types of underlying spot assets, which are the S&P 500 Index and the SPDR ETF. Two corresponding results of the frequency and magnitude of mispricing are shown, compared, and discussed. We observe a significant amount of mispricings, even in presence of the transaction cost, for both choices of underlying assets. But different patterns and distributions of mispricing are found. Furthermore, we gather the futures and spot trading data from different time periods to examine the impact of volatility on mispricing. Mispricing is more frequent in high volatility months. Finally, a multiple regression analysis is performed to study the effect of time-to-maturity, futures trading volume, dividend yield, and direction of mispricing on the absolute magnitude of mispricing. All the explanatory variables show a significant correlation with the magnitude of mispricing.

Index Spot-Futures Arbitrage: Evidence from Examining S&P 500 Index and SPDR ETF

By

Jianhao Zhang

Caroline Fohlin
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2021

Contents

List of Tables

List of Figures

Index Spot-Futures Arbitrage: Evidence from Examining S&P 500 Index and SPDR ETF

Jianhao Zhang

April 2021

1 Introduction

Index futures is a type of futures contract in which investors can buy or sell financial indices before a particular expiration date. The first stock index futures, based on the Value Line Composite Index, was issued in February 1982. Since then, the number of index futures has grown rapidly. Its enormous popularity benefits from its highly leveraged return, high liquidity, low commission cost, and the diversified portfolio associating with low risk. Similar to the commodity futures, the price of the futures contract reflects the futures value of the index, and the theoretical futures price for the cash index can be derived from the cost of carry model after adjusting for the risk-free interest rate and dividend yield (Paya and Peel, 2011; Ahn et al., 2002). In an efficient and competitive market, the asset is always traded at its fair value on exchanges, making it impossible for investors to purchase it for undervalued prices or sell it for inflated prices. Any price deviation from the theoretical fair value predicted by the model would be arbitraged away immediately in a frictionless market as claimed by Zhou (2017). However, many previous studies have identified the existence of persistent mispricing between the index futures and the associating spot asset. Yadav and

Pope (1994) show that index arbitrage opportunities are exploitable based on the empirical evidence from four years of synchronous hourly data from the UK futures market. More evidence of this mispricing relationship is discovered in other global futures markets, such as the ISE-30 in Turkey (McMillan and Ulku, 2009) and DAX index futures in German (Buhler and Kempf, 1995). In this context, a price discrepancy between the theoretical and actual value of the index futures and the corresponding arbitrage opportunity may be found.

This paper aims to examine the mispricing or arbitrage opportunities between the spot-futures relationship of the S&P 500 index (SPX) in the US market. We use minute-by-minute intraday E-mini S&P 500 Futures (ES) trading data to explore the frequency and magnitude of mispricing with respect to the underlying cash index. Two types of instruments are adopted as underlying cash assets in this study. The first one is to own a basket of stocks in specific weights in line with the SPX, since we cannot directly trade the index. But considering the amount of trading costs and the complexity of constructing the portfolio for 500 stocks, using such a spot asset to engage in the arbitrage activity can be very expensive and time-consuming. Thus, the introduction of index ETF (Exchange-Traded Fund) can be an ideal alternative to track and mirror the index. The ease of trading ETF, lack of short selling restriction, high liquidity, and low transaction cost provide competitive advantages compared to executing individual stock trades. But Frino and Gallagher (2002) point out the infeasibility of replicating the target index due to market frictions. The presence of tracking error, which measures the deviation of the ETF from its benchmark index, informs that two underlying spot assets are not the same. To deal with this issue, the SPDR S&P 500 ETF (SPY), which almost resembles the SPX exactly with minimal tracking error, is used as a tradable single underlying asset in this paper. By simultaneously examining two underlying assets, two corresponding results of the frequency and magnitude of mispricing are shown, compared, and discussed. Furthermore, we gather the futures and spot trading data from different time periods to examine the impact of volatility on mispricing. The different mispricing behavior is then analyzed for these volatility subsamples. Finally, a multiple regression analysis is performed to study the effect of time-to-maturity, futures trading volume, future and spot high-low spread, and direction of mispricing on the absolute magnitude of mispricing. The steps are then repeated in presence of the different sizes of transaction costs.

Our finding suggests that a significant amount of mispricings, even in presence of the transaction cost, is observed for both choices of underlying cash assets. But different patterns and distributions of mispricing are found, with negative mispricings dominating in the SPY model and positive mispricings dominating in the SPX model. High volatility leads to more frequent mispricings. The magnitude of mispricing behaves differently considering the change of transaction costs for two choices of underlying assets. Finally, all the explanatory variables are significantly correlated to the magnitude of mispricing at all representative quantiles.

The paper proceeds as follows. Section 2 discusses the literature that relates and contributes to this study. Section 3 describes the data as well as the selection and matching procedure. Section 4 introduces the models to estimate the theoretical futures price for two underlying assets and the regression model to examine the effect of the explanatory variables. The hypotheses of this research are also stated. Section 5 elaborates on interpreting the results. Section 6 concludes the paper.

2 Literature Review

2.1 Evidence of Mispricing

Much attention in the extant literature is dedicated to identifying and analyzing the mispricing or arbitrage opportunity between the index futures and cash market. Kurov and Lasser (2002) examine the effect of NASDAQ-100 index tracking stock (CUBES) on the price relationship between the NASDAQ-100 futures and the underlying index. They conclude that the average magnitude and frequency of mispricing fell after the introduction of

CUBES. In this study, however, a comparative analysis of mispricing frequency and magnitude is employed for SPY and SPX as two different underlying assets. Cummings and Frino (2011) examines the mispricing of Australian stock index futures. They found out that price volatility has a positive impact on the mispricing spread. We further contribute to study the effect of price volatility on the frequency of mispricing.

Richie et al. (2008) examine the index spot-futures mispricing by using both the SPDR ETF and the S&P 500 cash index as the underlying cash asset. They find evidence of persistent mispricing in both choices of underlying assets, with more negative mispricing appeared using the SPDR ETF and more positive mispricing presented using the cash index. The effect of volatility is also analyzed with greater market volatility leading to a higher frequency of mispricing. To add feasibility for engaging in the arbitrage activity, they explore the length of mispricing and the volume sufficiency for exploiting these opportunities. They concluded that the limits of arbitrage are due to the staleness of the cash index and the illiquidity of SPDR ETF.

This paper mainly employs and inherits the concept and approaches from Richie et al. (2008), but several contributions have been made to further add novelty to this area of study. First, the period of sample data has been extended from 2008 to 2020, in which the different market movements and mispricing behavior can be expected. Specifically, some major financial crises have occurred during this time frame, and it is worth exploring whether the mispricing relationship changed dramatically. We use multiple regression analysis to study how the absolute magnitude of mispricing gets impacted, committing to provide some perspectives for investors to establish their arbitrage positions.

2.2 Transaction Costs

In reality, the arbitrage opportunity is usually impeded by the trading costs. Even the mispricing between index futures and spot exists, the size of transaction costs determines the profitability of arbitrage opportunities. If transaction cost exceeds the magnitude of mispricing, arbitragers would indeed incur losses by exploiting the opportunity. Since this study does not assume a frictionless market, such a factor must be incorporated into the model to carry out the result. Therefore, a boundary condition is introduced to identify potential arbitrage opportunities. Within this bound, the arbitrage trades would not be induced. The size of transaction costs then determines the width of this no-arbitrage bound. Prior literature applies various means to estimate the transaction costs. Fremault (1991) uses a quadratic function to model the transaction cost. He particularly concerns about the symmetric costs, the round-trip costs when trading the assets, as well as the costs of immediacy, which investors pay to reduce execution lags. Maniar et al. (2010) adopt a constant to represent the transaction cost, which comprises of exchange fee, impact costs, and tax. Yadav and Pope (1990) express the percentage transaction costs as the summation of the percentage round-trip costs for equities, taxes in the percentage of asset value, percentage commissions in futures market, and percentage market impact cost. In our approach, the transaction costs are estimated as the percentage of theoretical futures price carrying out by the cost-of-carry model for analytical convenience. The mispricing behaviors under different transaction costs are compared and discussed.

2.3 Effect of Time-to-Maturity

Many prior works apply regression analysis to explore the relationship between the magnitude of mispricing and time-to-maturity. Buhler and Kempf (1995) look at the German index futures market and conclude that the number of arbitrage opportunities generates diverge patterns for futures near to the expiration date and futures far from the time to maturity. They find that arbitrageurs mainly take the opportunities when the futures are nearest to deliver. In addition, a significant positive relationship between the absolute magnitude of mispricing and time-to-maturity is observed (Yadav and Pope, 1990; MacKinlay and

Ramaswamy, 1988; Bhatt and Cakici, 1990). We also explore the impact of time-to-maturity, but we further extend our regression model to study the effect of future and spot high-low spread, index futures volume, and the direction of mispricing on the absolute magnitude of mispricing.

3 Data

The intraday trading records for S&P 500 Futures (ES), SPDR S&P 500 ETF (SPY), and S&P 500 Index (SPX) are collected from Firstrate Data. All three sets of data are using one-minute intervals. In order to examine the effect of the volatility on the mispricing behaviors, three months of data for each level of volatility are selected. The volatility level is determined by the S&P 500 index price fluctuation within a certain month. Specifically, for high volatility, 2008 Oct (Financial crisis), 2018 Dec, and 2020 March (Covid crisis) are included; For mid volatility, 2015 Aug, 2018 Feb, and 2020 Jun are selected; For Low volatility 2014 May, 2016 Aug, and 2017 Jun are used.

Due to the fact of different trading hours for each asset in the market, the sample size for each set is different. To eliminate the problem of non-synchronous trading, the quotes that mismatch their counterparts in time are discarded. In this way, the paired futures and underlying asset price at each minute can be obtained to be carried into the model. To reduce bias and in compliance with the approach of Richie et al. (2008), these data eliminate any futures price changes greater than 2 percent and any SPY price change greater than $0.75, which immediately reverse in the next transaction. The idea is to avoid a relatively large price change potentially caused by buying and selling pressure. As a result, 128335 paired observations are obtained.

The close price at every minute is used to represent the corresponding ES, SPY, and SPX price. In addition, the risk-free interest rate is taken from the one-year daily treasury yield curve rate from the website of the U.S. Department of Treasury. The quarterly dividend yield for SPX and quarterly cash dividend payment for SPY are also collected.

4 Methodology

4.1 Pricing Models

To calculate the theoretical futures price (TFP) for SPX, the standard cost of carry model for index futures is exploited, which in line with model used in many previous works as described in Section 1.

$$\text{SPX Model : TFP}_t = 10X_t \exp\left(\frac{(r-d)(T-t)}{365}\right) \tag{1}$$

where,

X_t = underlying SPX price at time t

r = risk-free interest rate

T = expiration date of the futures contract

t = date and time of transaction

$T - t$ = maturity date

d = SPX quarterly dividend yield

Similarly, to calculate the TFP for SPY, the modified form of cost of carry model (Richie et al., 2008) is adopted.

$$\text{SPY Model : TFP}_t = 10S_t \exp\left(\frac{r(T-t)}{365}\right) - 10D\left(\frac{T_{ds}}{91}\right) \tag{2}$$

where,

S_t = underlying SPY price traded at time t

r = risk-free interest rate

T = expiration date of the futures contract

t = date and time of transaction

$T - t$ = maturity date

D = SPY dividend cash payment per share

T_{ds}= number of days of holding SPY before the next dividend payment date

4.2 Frequency of Mispricing

First, it is important to define the transaction costs. If the costs exceed the price divergence between the theoretical futures price of the underlying asset and the actual futures price, the arbitrage opportunity is not profitable and should not be exploited. But to simplify the costs of arbitrage in the model, the total cost is estimated as a percentage of the theoretical futures price predicted by the above models. In this paper, transaction cost of 0.00%, 0.05%, 0.10%, 0.20%, 0.25%, and 0.50% of the theoretical futures price would be adopted. Arbitrage exists if:

$$|F_t - TFP_t| > C * TFP_t \tag{3}$$

where TFP_t is calculated from two models above, F_t is the futures contract price traded at time t, and C is size of transaction cost. And a no-arbitrage boundary would then be:

$$(1 - C)TFP_t < F_t < (1 + C)TFP_t \tag{4}$$

$$F_t > (1 - C)TFP_t = \text{lower boundary} \tag{5}$$

$$F_t < (1 + C)TFP_t = \text{upper boundary} \tag{6}$$

If the actual futures price is below the lower boundary, there exists negative mispricing or the futures is underpriced, and a minus signal is assigned to indicate the direction of mispricing. Under this situation, traders may long the futures and short the underlying cash asset. If the actual futures price is above the upper boundary, there exists positive mispricing or futures is overpriced, and plus signal is assigned. In reverse, traders may short the futures and long the underlying cash asset. Finally, if the actual futures price is between the lower and upper boundary, the mispricing condition no longer holds, and a zero signal will be assigned. No arbitrage position would be established in this case. Both models are examined by this boundary violation condition under various sizes of costs to identify the mispricing frequency. Intuitively, the width of the bound increases with a larger size of the transaction cost. Moreover, the volatility subsamples are also tested but only under costs of $0.00\%, 0.10\%$ and 0.20%.

4.3 magnitude of mispricing

According to the illustration of the boundary condition above, the magnitude of mispricing or the arbitrage profit is calculated by the equations below:

$$MS_{+} = F_t - (1 + C)TFP_t \tag{7}$$

$$MS_{-} = F_t - (1 - C)TFP_t \tag{8}$$

the magnitude of mispricing is calculated differently depending on the direction of mispricing. Observations with zero signal are excluded in this part. Intuitively, positive mispricing leads to positive magnitude and negative mispricing leads to negative magnitude. The sign of the overall magnitude of mispricing largely depends on the frequency of each signal. But to eliminate the effect of the direction of mispricing, the absolute magnitude of mispricing is calculated as:

$$MS_{abs} = |F_t - TFP_t| - C * TFP_t \tag{9}$$

We include the result of mean, median, standard deviation, min and max for magnitude of mispricing with or without sign. Results are also shown for different transaction costs and volatility subsamples.

4.4 Regression Analysis

It is interesting to see whether the mispricing is random or whether it can be explained by some variables. To accomplish the goal, the absolute magnitude of mispricing is regressed against time-to-maturity, futures trading volume, future and spot high-low spread, and the sign of mispricing. In particular, we use a dummy variable for the direction of mispricing, 0 indicating negative mispricing and 1 indicating positive mispricing. The high-low spread is a ratio calculated by:

$$HL_{Spread} = \frac{High - Low}{Close} \tag{10}$$

The high, low, and close are the corresponding prices for each observation at the minute. We use this as a measure of liquidity in this paper to see how the absolute magnitude of mispricing gets impacted by it.

We apply a quantile regression model because the absolute magnitude of mispricing is not normally distributed in both models, and a substantial amount of outliers exist within the sample. This approach has several advantages. First, the model is robust to outliers. Second, the regression model flexes the distributional assumption and fits heterogeneous data. We then are able to explore the value of the coefficient at different quantiles of the dependent variables. The regression model is:

$$MS_{abs}(\tau) = \beta_1(\tau)T + \beta_2(\tau)FL + \beta_3(\tau)SL + \beta_4(\tau)V + \beta_5(\tau)S \tag{11}$$

where T represents the time-to-maturity for futures contracts, FL is the futures high-low spread, SL is the spot high-low spread, V indicates the futures trading volume, and S represent the direction of mispricing. τ represents the coefficient at indicated quantile of absolute magnitude of mispricing.

4.5 Hypothesis

Several hypotheses concerning the mispricing behaviors would be tested in the next section:

Hypothesis I: There are no arbitrage opportunities after considering the transaction costs.

Hypothesis II: Both SPY and SPX model should generate similar patterns and distributions for the frequency of mispricing.

Hypothesis III: The market volatility has no impact on the number of mispricings.

Hypothesis IV: The absolute mean magnitudes of mispricing for two models should react the same to the size of transaction cost.

Hypothesis V: The market volatility has no impact on the magnitude of mispricings.

Hypothesis VI: There is no significant difference between the magnitude of mispricing for SPX and SPY model.

5 Results

5.1 Frequency of Mispricing

Table 1 shows the result of the frequency of mispricing for both SPY and SPX as the underlying cash asset for a total sample of 123250 observations. The representative transaction costs are varied in six levels, ranging from 0.00% to 0.50%. In addition, the signal is assigned to indicate the direction of mispricing for both models with minus signal (negative mispricing) representing the violation of lower boundary, zero signal (true pricing) referring to the no-arbitrage condition, and plus signal (positive mispricing) suggesting the violation of upper boundary. Intuitively, in presence of the transaction cost, the frequency of mispricing drops significantly because more observations would lay within the lower and upper boundary. This phenomenon stands for both positive and negative mispricings in two models. But still, the significant amount of boundary violations for both models suggests the arbitrage opportunity exists no matter which instrument is adopted as the underlying cash asset, which does not support Hypothesis I that no arbitrage opportunity would be observed in presence of the transaction cost.

However, these two models still behave differently in several ways. First, the distribution of frequency of mispricing differs in the two models. At transaction cost of 0.00%, the SPY model has 64.53% of negative mispricing compared to 35.47% of positive mispricing while the SPX model indicates a distribution of 43.91% of negative mispricing and 56.09% of positive mispricing. Therefore, there is a propensity for the actual futures price to fall below the theoretical fair price when using the SPY as the underlying cash asset. This pattern of mispricing proportion appears at all sizes of transaction costs. Since the violation of the lower boundary implies the futures is undervalued, the investors should exploit the arbitrage opportunity by longing the futures and shorting the SPY. On the other hand, the investors should seek the arbitrage profit of shorting the futures and longing the stocks when SPX becomes the underlying cash asset. The strategy of SPX model aligns with the notion that a long position in stocks is more feasible than a short position, taking the short selling restriction into account.

Second, the distinct frequency of zero signal also proves the difference between the two models. At all representative sizes of transaction costs, the frequency of zero signal in SPX model exceeds that of in SPY model by a huge amount. This notion demonstrates that the S&P 500 futures contract is inclined to be priced correctly when employing the cash index as the underlying asset. In addition, note that 52.22% of observations in SPX model do not violate the boundary condition at transaction cost of 0.05%, while 52.80% of observations in SPY model lay within the no-arbitrage boundary at a transaction cost of 0.20%. Therefore, as the size of transaction costs increases, the amount of mispricing diminishes faster when exploiting SPX as the underlying asset. At 0.50% transaction cost, there are more than double violations in the SPY model compared to the SPX model, which

Table 1: Frequency of Mispricing for Entire Sample

TC	Total obs	model 1: SPY			model 2: SPX		
		-1	0	1	-1	0	1
0.00%	123250	79530	0	43720	54123	0	69127
		64.53%	0.00%	35.47%	43.91%	0.00%	56.09%
0.05%	123250	69489	16091	37670	19995	64364	38891
		56.38%	13.06%	30.56%	16.22%	52.22%	31.55%
0.10%	123250	55903	34034	33313	5052	96710	21488
		45.36%	27.61%	27.03%	4.10%	78.47%	17.43%
0.20%	123250	34617	65070	23563	467	110982	11801
		28.09%	52.80%	19.12%	0.38%	90.05%	9.57%
0.25%	123250	30052	77920	15278	282	112414	10554
		24.38%	63.22%	12.40%	0.23%	91.21%	8.56%
0.50%	123250	14328	108873	49	108	116718	6424
		11.63%	88.34%	0.04%	0.09%	94.70%	5.21%

suggests that investors should be disposed to trade SPY to exploit arbitrage opportunity. In this case, Hypothesis II is violated by the distinctive pattern and distribution for the frequency of mispricing in the two models.

Table 2 examines the impact of market volatility on the frequency of mispricing. The entire sample is filtered into three subsamples by the level of market volatility to investigate whether the volatility affects the distribution and pattern of mispricing. The SPY model, in consistent with the result in Table 1, shows that the negative mispricing dominates the positive mispricing for each volatility sample. The SPX model reflects that the majority of mispricing is positive, excluding the mid volatility at a transaction cost of 0.00%, which also conforms with the previous result. In addition, the number of mispricing tends to decrease as the level of volatility declines from high to low for both models at all sizes of transaction costs, which violates Hypothesis III that the market volatility has no impact on the number

Table 2: Frequency of Mispricing for Volatility Subsamples

TC	vol	Total obs	model 1: SPY			model 2: SPX		
			-1	0	1	-1	0	1
0.00%	high	38537	24021	0	14516	12237	0	26300
			62.33%	0.00%	37.67%	31.75%	0.00%	68.25%
	mid	41415	28441	0	12974	24186	0	17229
			68.67%	0.00%	31.33%	58.40%	0.00%	41.60%
	low	43298	27068	0	16230	17700	0	25598
			62.52%	0.00%	37.48%	40.88%	0.00%	59.12%
0.10%	high	38537	21249	4081	13207	2341	21450	14746
			55.14%	10.59%	34.27%	6.07%	55.66%	38.26%
	mid	41415	16385	12790	12240	2230	36256	2929
			39.56%	30.88%	29.55%	5.38%	87.54%	7.07%
	low	43298	18269	17163	7866	481	39004	3813
			42.19%	39.64%	18.17%	1.11%	90.08%	8.81%
0.20%	high	38537	12307	14579	11651	347	27154	11036
			31.94%	37.83%	30.23%	0.90%	70.46%	28.64%
	mid	41415	9415	24758	7242	119	40740	556
			22.73%	59.78%	17.49%	0.29%	98.37%	1.34%
	low	43298	12895	25733	4670	1	43088	209
			29.78%	59.43%	10.79%	0.00%	99.51%	0.48%

of mispricing. This pattern implies that more arbitrage opportunities would exist in a higher volatile month, agreeing with Draper and Fung (2003) conclusion of a positive correlation between the futures mispricing and market volatility. Moreover, Cummings and Frino (2011) proposed that spot and futures prices respond to the market at a different speed, and futures contracts are more efficient in delivering and processing the new information (Rossi and Santucci de Magistris, 2013), which results in the index futures price deviating from the cost of carrying the index basket of stocks. Therefore, in the volatile period, when both futures and spot asset prices are highly fluctuated and quickly changing, this effect is even magnified, and a relatively large number of mispricing can be explained.

5.2 Magnitude of Mispricing

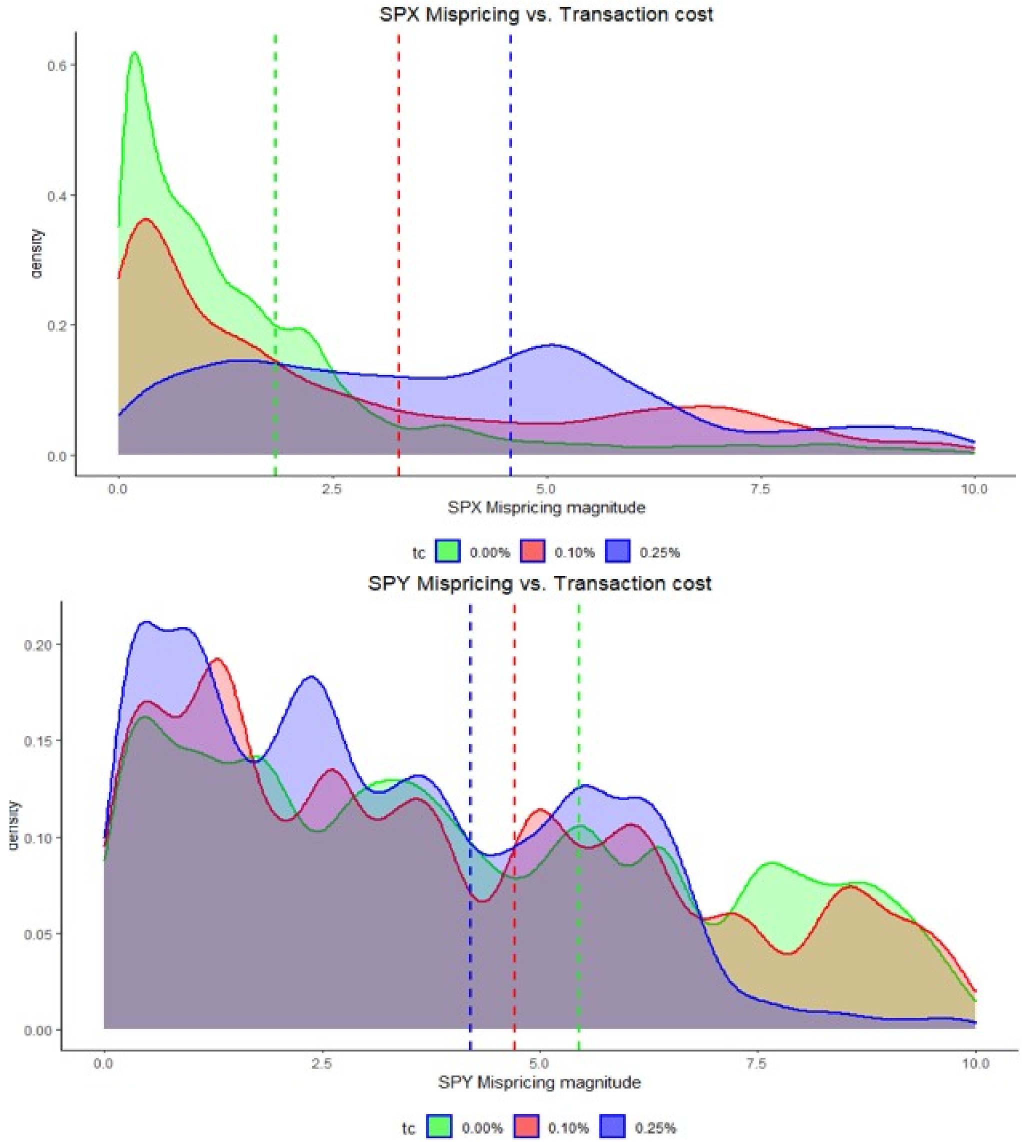
Table 3 lists the summary statistics of mispricing magnitude for both models at all representative sizes of transaction costs. At each level of transaction cost, the first row is the magnitude of mispricing, considering the direction of violations, and the second row is the corresponding absolute value basis. As for the SPY model, the absolute mean of mispricing drops from 5.42 at a cost of 0.00% to 3.03 at a cost of 0.50%. This decreasing trend is also true for the median of mispricing. The simple and direct implication is that larger transaction cost is associated with lower arbitrage profit. However, a similar pattern is not observed in the SPX model. Instead, the average mispricing magnitude increases with larger transaction costs, rejecting Hypothesis IV that indicates the absolute mean magnitudes of mispricing in two models behave in the same way when including the transaction costs. This counter-intuitive relationship suggests that the investor would gain more arbitrage profit when they actually lose more money due to rising costs.

To explore the reason for such an unparallel result between two models, the density curve for Magnitude of Mispricing vs. Transaction cost is plotted for both models in Figure 1. The density plot for SPX model is skewed to the right at 0.00% transaction cost, indicating that the mispricing is heavily dense around the magnitude of 0.00, which implies a large proportion of small magnitude existing within the sample. This phenomenon contributes to a small sample mean even at a low level of the transaction cost. However, when the transaction cost is elevated, the proportion of small magnitude is dropped significantly because many observations are no longer mispriced according to the boundary condition. These observations are not taken into consideration when calculating the sample mean. As evidenced by the curve at the 0.50% transaction cost, the majority of the mispricing magnitude becomes centered around 5.0. Therefore, the increased proportion of large magnitude and a smaller sample of mispriced observations lead to a contradictory large mean. On the other hand, the density of mispricing in the SPY model is more evenly distributed with a

Table 3: Magnitude of Mispricing for Entire Sample

TC	SPY model						SPX model					
	N	mean	median	sd	min	max	N	mean	median	sd	min	max
0.00%	123250	-2.02	-1.68	6.72	-58.62	104.22	123250	0.89	0.19	2.92	-62.69	46.45
		5.42	4.42	4.46	0.00	104.22		1.83	1.06	2.44	0.00	62.69
0.05%	107159	-1.97	-1.38	6.27	-57.42	103.02	58886	1.55	0.59	3.37	-61.47	45.27
		4.99	4.03	4.28	0.00	103.02		2.20	1.02	2.99	0.00	61.47
0.10%	89216	-2.04	-1.12	5.92	-56.22	101.82	26540	2.82	1.56	3.99	-60.25	44.08
		4.70	3.73	4.13	0.00	101.82		3.26	1.73	3.64	0.00	60.25
0.20%	58180	-2.67	-1.73	5.14	-53.83	99.42	12268	4.33	4.26	4.08	-57.81	41.70
		4.28	3.31	3.91	0.00	99.42		4.69	4.38	3.66	0.00	57.81
0.25%	45330	-3.24	-2.93	4.65	-52.63	98.23	10836	4.25	4.10	3.97	-56.60	40.52
		4.17	3.22	3.83	0.00	98.23		4.59	4.20	3.57	0.00	56.60
0.50%	14377	-3.02	-1.44	3.49	-46.64	92.23	6532	3.50	2.85	3.55	-50.50	35.71
		3.03	1.44	3.48	0.00	92.23		3.80	2.88	3.23	0.00	50.50

Figure 1: Density Plot for Mispricing vs. Transaction cost



significant amount of mispricing appears at all levels of magnitude. Therefore, when the transaction cost is increased, the mispricing magnitude still shrinks but only a relatively smaller number of observations are excluded from the sample. This stable sample causes the curve to be shaped in a similar pattern. In the Figure, only the proportion of extreme mispricing with a magnitude between 7.50 and 10.00 is reduced dramatically, leading the rest to have a higher density. In this case, the absolute mean is negatively related to the transaction cost.

The above description demonstrates that the theoretical futures price is closer to the actual futures price when using SPX as the underlying spot asset, as evidenced by the large proportion of mispricing magnitude extremely close to zero. Although the size of arbitrage profit for SPX model becomes larger at a higher transaction cost level, the investors would not be better off to seek the arbitrage opportunity because the amount of mispricing also diminishes. Even more, those arbitrage trades are hard to accomplish because the extreme mispricing magnitude is usually associated with limited trading volume and price staleness, while more opportunities and more logical distribution of the profit can be explored by the arbitragers to use the SPY as the underlying asset.

Table 4 further elaborates on the summary statistics in Table 3 to study the impact of volatility on the magnitude of mispricing. In congruent with the previous result, at each level of volatility, a decreasing trend of average mispricing magnitude is observed as the transaction cost elevated in the SPY model instead of the SPX model. However, when the level of transaction cost is controlled, a positive relationship between the mean mispricing and the volatility is only presented in the SPX model, rejecting Hypothesis V. Additionally, the absolute mean mispricing for SPY model is higher than SPX model at all levels of volatility, most likely because of the higher cost to generate a portfolio equivalent to the actual index.

Table 5 and Table 6 examines the difference between the magnitude of mispricing in the two models. The two-sample t-test is performed and results are listed in the last column in both tables. In Table 5, the absolute mean difference of mispricings diminishes with higher transaction cost, indicating that two models have mispricings closer to each other. The positive absolute mean difference conforms with the previous conclusion that a larger mispricing magnitude is found in the SPY model. Furthermore, the large t-statistic values, excluding the one at 0.50% transaction cost, show that the difference in the mean difference between two models is statistically significant at 1.00% level, which violates hypothesis VI,

Table 4: Magnitude of Mispricing for Volatility Subsamples

TC	vol	model 1: SPY						model 2: SPX					
		N	mean	median	sd	min	max	N	mean	median	sd	min	max
0.00%	high	38537	-2.09	-2.28	7.76	-58.62	104.22	38537	2.14	0.75	4.27	-62.69	46.45
			6.27	5.29	5.03	0.00	104.22		3.17	1.72	3.58	0.00	62.69
	mid	41415	-1.53	-1.76	6.80	-42.36	12.11	41415	0.13	-0.33	2.06	-57.00	27.66
			5.54	4.50	4.23	0.00	42.36		1.44	1.17	1.48	0.00	57.00
	low	43298	-2.41	-1.41	5.52	-41.04	8.96	43298	0.50	0.18	1.36	-4.42	10.07
			4.54	3.18	3.95	0.00	41.04		1.02	0.68	1.03	0.00	10.07
0.10%	high	34456	-1.71	-1.05	6.39	-56.22	101.82	17087	3.98	3.35	4.26	-60.25	44.08
			4.70	3.46	4.65	0.00	101.82		4.42	3.52	3.80	0.00	60.25
	mid	28625	-1.46	-0.32	6.03	-40.49	9.00	5159	0.58	0.14	2.92	-55.10	25.74
			4.78	3.62	3.95	0.00	40.49		1.32	0.52	2.66	0.00	55.10
	low	26135	-3.11	-2.18	4.92	-38.86	6.54	4294	0.91	0.82	0.93	-2.26	7.64
			4.62	4.68	3.54	0.00	38.86		0.96	0.82	0.88	0.00	7.64
0.20%	high	23958	-1.99	-0.06	5.75	-53.83	99.42	11383	4.57	4.54	3.86	-57.81	41.70
			4.00	2.29	4.59	0.00	99.42		4.82	4.62	3.54	0.00	57.81
	mid	16657	-2.68	-0.87	5.36	-38.63	5.90	675	1.35	1.82	6.11	-53.21	23.83
			4.73	4.28	3.68	0.00	38.63		3.64	2.36	5.09	0.00	53.21
	low	17565	-3.59	-4.00	3.72	-36.69	4.12	210	0.93	0.68	0.90	-0.10	5.20
			4.24	4.02	2.94	0.00	36.69		0.93	0.68	0.90	0.00	5.20

Table 5: Difference in Magnitude of Mispricing for Entire Sample

TC	SPY Mispricing - SPX Mispricing						
	N	mean	median	sd	min	max	t-statisitic
0.00%	123250	-2.91	-3.21	7.60	-52.85	80.13	-139.16
		3.58	2.72	5.05	50.95	80.13	247.21
0.05%	52886	-5.00	-3.58	6.62	-50.45	80.17	-169.63
		2.47	1.26	5.60	50.93	80.17	98.82
0.10%	21168	-6.95	-6.03	6.51	-48.05	80.21	-138.14
		1.95	0.70	6.67	50.90	80.21	44.34
0.20%	7731	-7.05	-6.02	4.85	-43.25	80.29	-95.53
		0.23	1.59	6.27	50.85	80.29	-3.61
0.25%	6022	-5.85	-5.06	4.33	-40.85	80.33	-77.49
		0.66	1.60	5.57	50.82	80.33	-9.71
0.50%	220	-2.56	-2.45	9.38	-23.58	80.53	-2.98
		0.49	0.18	9.01	38.15	80.53	-0.62

and two samples are truly different from each other. Table 6 replicates the steps on the volatility subsample and gives a similar result.

5.3 Multiple Regression Analysis of the Magnitude of Mispricing

To further explore the factors that influence the magnitude of mispricing, a multiple regression analysis is performed for both models. The time-to-maturity, dividend payment or yield, futures volume, and direction of mispricing are used as independent variables. This analysis can provide some perspectives on whether the mispricing is random or correlated to these variables. According to Figure 1, the absolute magnitude of mispricing is not normally distributed in both models, and a substantial amount of outliers exist within the sample. Since the key assumptions of OLS linear regression cannot be met, the quantile regression is adopted to study the relationship between conditional quantile of the magnitude of mispricing

Table 6: Difference in Magnitude of Mispricing for Volatility Subsamples

TC	vol	SPY Mispricing - SPX Mispricing						
		N	mean	median	sd	min	max	t statistic
0.00%	high	38537	-4.23	-4.43	8.75	-52.85	80.13	-93.83
			3.10	3.49	6.58	50.95	80.13	98.48
	mid	41415	-1.66	-1.66	7.53	-44.67	48.34	-47.56
			4.10	2.91	4.53	48.34	34.41	185.97
	low	43298	-2.91	-0.53	6.25	-42.95	8.69	-106.69
			3.52	1.85	3.73	4.95	39.12	179.30
0.10%	high	13854	-6.88	-5.75	6.21	-48.05	80.21	-110.19
			0.28	0.53	6.71	50.90	80.21	5.25
	mid	3627	-5.77	-1.78	8.23	-39.84	48.29	-46.10
			3.93	0.60	6.13	48.29	34.38	33.23
	low	3687	-8.34	-9.69	5.30	-35.30	6.10	-107.33
			6.29	7.52	4.03	0.94	30.90	81.15
0.20%	high	6950	-6.48	-5.75	4.00	-43.25	80.29	-94.65
			1.14	2.86	5.48	50.85	80.29	-19.23
	mid	575	-11.86	-9.08	8.60	-35.68	38.26	-33.60
			6.82	5.99	7.47	38.26	31.55	21.29
	low	206	-12.63	-9.06	5.72	-30.40	-5.75	-35.00
			10.74	8.56	4.57	2.01	22.92	29.77

and the explanatory variables. This approach has several advantages. First, the model is robust to outliers. Second, the regression model flexes the distributional assumption and fits heterogeneous data. We then are able to explore the value of the coefficient at different quantiles of the dependent variables.

The regression result in Table 7 reveals that the time-to-maturity positively and significantly affects the magnitude of SPY mispricing at all representative quantiles, agreeing with findings of Yadav and Pope (1994). Theoretically, the actual futures price converges to the theoretical futures price when approaching the expiration date. Therefore, this positive relationship can be testified because far contract leads to a wider gap and divergence between the actual and theoretical futures price so that a greater mispricing magnitude can

Table 7: Regression Result for SPY Model

	Dependent variable:				
	SPY Mispricing				
	tau=0.1	tau=0.25	tau=0.5	tau=0.75	tau=0.9
	(1)	(2)	(3)	(4)	(5)
Time-to-Maturity	0.024***	0.042***	0.068***	0.085***	0.106***
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0002)
Future High-Low Spread	-17.296***	30.654***	60.384***	306.172***	450.655***
	(3.562)	(4.460)	(3.038)	(17.493)	(10.283)
SPY High-Low Spread	-0.00004***	-0.0001***	-0.0002***	-0.0001***	-0.0001***
	(0.00000)	(0.00000)	(0.00000)	(0.00001)	(0.00000)
Future Volume	0.125***	1.288***	4.796***	7.764***	9.931***
	(0.007)	(0.016)	(0.055)	(0.021)	(0.032)
Signal of Mispricing	-0.331***	-0.798***	-1.311***	-1.619***	-2.229***
	(0.007)	(0.007)	(0.021)	(0.014)	(0.008)
Observations	123,250	123,250	123,250	123,250	123,250
Note:					*p<0.1; **p<0.05; ***p<0.01

be explained. At higher quantile, the effect of time-to-maturity is magnified. The coefficient of future high-low spread is also significant and positive except at 0.1 quantile, which states that a larger price spread leads to a larger magnitude of mispricing. The coefficient value increases dramatically at higher quantiles, suggesting the absolute magnitude of mispricing would be largely impacted by the future high-low spread at top quantiles. The coefficient of SPY high-low spread is negative, significant, and relatively stable at each quantile. Moreover, the extremely small coefficient suggests that the SPY high-low spread is unlikely to have an impact on the absolute magnitude of mispricing. The coefficient of futures trading volume

is positive and significant. Moreover, the positive effect of future volume is large for the

Table 8: Regression Result for SPY Model

	Dependent variable:				
	SPX Mispricing				
	tau=0.1	tau=0.25	tau=0.5	tau=0.75	tau=0.9
	(1)	(2)	(3)	(4)	(5)
Time-to-Maturity	0.002***	0.006***	0.013***	0.023***	0.035***
	(0.00004)	(0.0001)	(0.0001)	(0.0001)	(0.0004)
Future High-Low Spread	51.534***	117.411***	289.101***	609.848***	1,064.524***
	(5.803)	(9.401)	(11.278)	(21.142)	(41.930)
SPX High-Low Spread	9.935	63.790***	109.989***	350.509***	742.954***
	(7.418)	(12.287)	(13.773)	(28.497)	(57.047)
Future Volume	0.00001***	0.00001***	0.00002***	0.00001***	0.00002***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Signal of Mispricing	0.028***	0.031***	0.017***	-0.070***	0.114***
	(0.003)	(0.004)	(0.006)	(0.004)	(0.032)
Observations	123,250	123,250	123,250	123,250	123,250
Note:					*p<0.1; **p<0.05; ***p<0.01

magnitude of mispricing at the top quantiles. Finally, a negative and significant coefficient for the signal of mispricing illustrates that positive mispricings reduce the magnitude. The negative effect tends to be larger at higher quantiles. These results stay robustly in presence of the transaction cost.

Table 8 presents the regression results for the mispricing in SPX model, in which the regression result for time-to-maturity is similar. The coefficient of future high-low spread is also positive and significant in SPX model, but the coefficient is larger comparing to that in SPY model at each quantile. In addition, the coefficient of SPX high-low spread is positive and significant, which also increases at higher quantiles. In the SPX model, the spot price spread does have a large effect on the magnitude of mispricing. The coefficient for future volume still indicates a positive relationship with the magnitude of mispricing, but its effect can almost be neglected. The effect of the signal of mispricing also behaves in a distinctive pattern, which is significant and positive in this case, excluding the one at 0.75 quantile. This change demonstrates that positive mispricing has a larger magnitude of mispricing relative to negative mispricing. These results are also robust in presence of transaction costs.

6 Conclusion

In this paper, two different underlying assets are chosen to explore the spot-futures arbitrage relationship for S&P 500. The findings show that a substantial amount of mispricings is observed when either SPY or SPX is employed as the underlying cash asset. This result stays robustly even in presence of transaction costs. A comparative analysis is also made to study the difference in choices of the underlying cash assets. In particular, more negative mispricings are found using SPY, and more positive mispricing mispricings are found using SPX. The number of mispricings in SPY model exceeds that of in SPX model in all sizes of transaction cost. The absolute mean magnitude of mispricing behaves distinctively as the transaction costs increase, which can be explained by the different density proportions within each sample. And the magnitudes of mispricing for the two models are significantly different from each other. In addition, analysis on volatility subsamples shows that more frequent mispricings are observed in high volatility months. But only in SPX model do we observe the mean magnitude of mispricing decreases with the level of volatility. Finally, the quantile regression results state that the time-to-maturity, futures trading volume, future and spot high-low spread, and direction of mispricing are all significantly related to the magnitude of mispricing in both models. However, the results of our analysis could change when adopting different measures of the risk-free interest rate, dividend yield for underlying assets, and the realistic transaction cost in the market.

References

Ahn, D.-H., Boudoukh, J., and Richardson, M., a. W. R. F. (2002). Partial adjustment or stale prices? implications from stock index and futures return autocorrelations. *Review of Financial Studies*, 15(2):655–689.

Bhatt, S. and Cakici, N. (1990). Premiums on stock index futures–some evidence. *Journal of Futures Markets*, 10(4):367–375.

Buhler, W. and Kempf, A. (1995). Dax index futures: Mispricing and arbitrage in german markets. *Journal of Futures Markets*, 15(7):833–859.

Cummings, J. R. and Frino, A. (2011). Index arbitrage and the pricing relationship between australian stock index futures and their underlying shares. *Accounting Finance*, 51(3):661–683.

Draper, P. and Fung, J. K. W. (2003). Discretionary government intervention and the mispricing of index futures. *Journal of Futures Markets*, 23(12):1159–1189.

Fremault, A. (1991). Stock index futures and index arbitrage in a rational expectations model. *Journal of Business*, 64(4):523–547.

Frino, A. and Gallagher, D. R. (2002). Is index performance achievable? an analysis of australian equity index funds. *Abacus*, 38(2):200.

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

MacKinlay, A. and Ramaswamy, K. (1988). Index-futures arbitrage and the behavior of stock index futures prices. *Review of Financial Studies*, 1(2).

Maniar, H. M., Bhatt, R., and Maniyar, D. M. (2010). Price discovery and arbitrage between futures and cash markets: A case study on national stock exchange of india (nse). *Finance India*, 24(3):929–944.

McMillan, D. G. and Ulku, N. (2009). Persistent mispricing in a recently opened emerging index futures market: Arbitrageurs invited. *Journal of Futures Markets*, 29(3):218–243.

Paya, I. and Peel, D. A. (2011). Systematic sampling of nonlinear models: Evidence on speed of adjustment in index futures markets. *Journal of Futures Markets*, 31(2):192–203.

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

Rossi, E. and Santucci de Magistris, P. (2013). A no-arbitrage fractional cointegration model for futures and spot daily ranges. *Journal of Futures Markets*, 33(1):77–102.

Yadav, P. K. and Pope, P. F. (1990). Stock index futures arbitrage: International evidence. *Journal of Futures Markets*, 10(6):573–603.

Yadav, P. K. and Pope, P. F. (1994). Stock index futures mispricing: Profit opportunities or risk premia? *Journal of Banking and Finance*, 18(5):921–953.

Zhou, J. (2017). Index arbitrage and dynamics between reit index futures and spot prices. *Applied Economics*, 49(19-21):1875–1885.