

The Microstructure of Arbitrage: ETF Evidence

Ben R. Marshall
Massey University
B.Marshall@Massey.ac.nz

Nhut H. Nguyen*
University of Auckland
n.nguyen@auckland.ac.nz

Nuttawat Visaltanachoti
Massey University
N.Visaltanachoti@Massey.ac.nz

Abstract

We use two extremely liquid and close substitute S&P 500 ETFs to analyze the prevailing trading conditions when mispricing allowing arbitrage opportunities is created. Spreads increase just before arbitrage opportunities, consistent with a decrease in liquidity. Order imbalance increases as markets become more one-sided and spread changes become more volatile which suggest an increase in liquidity risk. Price deviations are economically significant and followed by a tendency to quickly correct back towards parity. Our findings complement the growing literature highlighting the importance of liquidity dynamics on the price discovery process in well-functioning financial markets.

JEL Classification: G1, G14

Keywords: Arbitrage, Pairs Trading, ETF

First Version: 23 September 2010

This Version: 29 February 2012

* Corresponding author: Nhut H. Nguyen, Department of Accounting and Finance, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand; telephone +64 9 3737599 ext. 83326; fax +64 9 3737406; e-mail n.nguyen@auckland.ac.nz. We thank Andrea Bennett, participants at the 2011 FMA Conference, and in particular, Dale WR Rosenthal (our session chair) and James Conover (our discussant), and participants at the 2011 New Zealand Finance Colloquium for valuable comments. All errors are our own.

The Microstructure of Arbitrage: ETF Evidence

Abstract

We use two extremely liquid and close substitute S&P 500 ETFs to analyze the prevailing trading conditions when mispricing allowing arbitrage opportunities is created. Spreads increase just before arbitrage opportunities, consistent with a decrease in liquidity. Order imbalance increases as markets become more one-sided and spread changes become more volatile which suggest an increase in liquidity risk. Price deviations are economically significant and followed by a tendency to quickly correct back towards parity. Our findings complement the growing literature highlighting the importance of liquidity dynamics on the price discovery process in well-functioning financial markets.

JEL Classification: G1, G14

Keywords: Arbitrage, Pairs Trading, ETF

1. Introduction

Arbitrage is frequently described as one of the most important principles in finance. Numerous papers report situations of mispricing which appear to allow for arbitrage profits to be made (e.g. Froot and Dabora, 1999; Mitchell, Pulvino, and Stafford, 2002; and Gagnon and Karolyi, 2010).¹ Others point out that arbitrage is rarely, if ever, as risk-free and costless as the text-book definition suggests (e.g. De Long, Shleifer, Summers, and Waldman, 1990; Shleifer and Vishny, 1997). Much less is known about the intraday market characteristics that exist when arbitrage opportunities are created. It is this “microstructure of arbitrage” strand of the literature that we contribute to.

Our evidence suggests that a fall in liquidity combined with an increase in liquidity risk contribute to the arbitrage opportunities. We show spreads increase prior to arbitrage opportunities, which is consistent with a decline in liquidity. Order imbalance increases, consistent with the market becoming more one-sided. The standard deviation of spread changes increases which is indicative of an increase in liquidity risk. Trade value also increases which, according to Johnson (2008), is evidence of an increase in liquidity volatility. Johnson (2008) explains the positive relation between liquidity risk and volume as follows (p. 411) “Intuitively, large changes in liquidity cannot occur without a lot of population flux, and a small amount of flux must imply a small change.” Lastly, the standard deviation of trade-to-trade returns increases.

¹ The strategies used to capture this mispricing are variously referred to as “convergence trading”, “pairs trading”, and “statistical arbitrage.” Papers applying these strategies include: Gatev, Goetzmann, and Rouwenhorst (2003), Engleberg, Gao, and Jagannathan (2010), Hogan, Jarrow, Teo, and Warachka (2004). Theoretical papers include Xiong (2001) and Bondarenko (2003).

Our paper is related to the important work of Schultz and Shive (2010). They show, using dual-class shares, that the most liquid share is the one most often responsible for creating arbitrage opportunities. Once mispricing occurs trading volume in the underpriced (overpriced) share shifts from sales (purchases) to purchases (sales) and trading becomes more urgent. Schultz and Shive (2010) also show one-sided trades, rather than long-short trades, are used to eliminate the majority of the mispricing. Our focus on the level of prevailing liquidity, liquidity risk and other security characteristics that exist when mispricing that is large enough for arbitrage opportunities occurs complements Schultz and Shive (2010)'s microstructure of arbitrage findings. Our findings are consistent with Deuskar and Johnson (2011) who show that unpredictable flow driven risk accounts for around half of market variance.

S&P 500 ETFs are the instruments used in our study. These ETFs are natural choices for an arbitrage microstructure study for a number of reasons. Firstly, they are highly liquid and widely followed. A large number of investors track these instruments and trade them (e.g. Elton, Gruber, Commer, and Li, 2002). The SPDR Trust (ticker SPY) is more liquid than any stock and the iShare (ticker IVV) is also highly liquid.² Any arbitrage opportunities can therefore be easily exploited. ETFs are particularly suitable instruments for identifying the role liquidity plays in mispricing. It is well accepted that the risk to uninformed investors of trading against an investor with private or asymmetric information is considerably lower for investments in ETFs rather than individual stocks.³ This means that intraday changes in spreads are relatively (compared to stocks) more likely to be driven by liquidity changes than asymmetric information.

² Vanguard listed an S&P 500 ETF in September 2010, but we do not include this in our analysis due to insufficient data. There are few other ETF pairs that track the same instrument. Where these do exist, the volume is typically heavily concentrated in one of the pairs, leaving relatively long intraday periods with no trading in the more illiquid security, which is problematic for a microstructure study such as this. An exception is the two ETFs on Gold.

³ See Hamm (2011) footnote 1.

Moreover, “convergence risk” is much lower than in most arbitrage settings. John Maynard Keynes once remarked “the market can stay irrational longer than you can stay solvent.”⁴ This creates a risk for arbitrageurs as simply identifying and trading on mispricing does not mean profits will be made quickly or at all. Different aspects of this arbitrage risk have been documented. De Long, Shleifer, Summers, and Waldman (1990), focus on the risk of further price divergence due to the actions of irrational “noise” traders. Abreu and Brunnermeier (2002, p. 343) highlight that arbitrageurs still face synchronization risk or “uncertainty regarding the timing of the price correction” even when noise traders are not present. Mitchell, Pulvino, and Stafford (2002) refer to “horizon risk”, and Shleifer and Vishny (1997) highlight the “margin risk” of leveraged positions having to be liquidated due to margin calls before the final convergence occurs.

Convergence risk is minimized in our market setting for a number of reasons. Firstly, the SPY and IVV compete for investor funds. Their ability to closely track the underlying index is an important aspect of this so the management of each fund have an incentive to minimize tracking error. Secondly, as Ackert and Tian (2010) note, institutional investors can exchange each ETF for the underlying stocks.⁵ Larger ETF price divergence creates more incentive for this activity.

Another risk of arbitrage is fundamental risk. This refers to the fact that the two assets are not perfect substitutes or identical in all respects. It is extremely difficult to find assets that are completely identical. Dual class shares are likely to have different voting rights (e.g. Schultz and Shive, 2010). Dual listed stocks in different countries are frequently subject to different institutional features such as liquidity differences and index inclusion in one country (e.g. Froot and Dabora, 1999). Short-selling constraints in one country may play a role and there may also

⁴ <http://www.maynardkeynes.org/keynes-the-speculator.html>

⁵ See also Engle and Sarkar (2006) and Aber, Li, and Can (2009) for more information.

be “tax-induced investor heterogeneity” (e.g. Froot and Dabora, 1999, p. 215). While minimized, fundamental risk does exist in our market setting. The SPY and IVV both have the aim of mirroring the S&P 500 index, but there may be small differences in the composition of the basket of securities each ETF uses to track the S&P 500. As noted in the SPDR prospectus⁶, the trustees are permitted to allow a certain amount of “misweighting” between each stock in the SPDR and the S&P 500 index each day. It is only when the misweighting in a stock exceeds 150% of this amount that a re-balancing is required. There may also be differences in the actual securities held by each fund. The IVV fund documentation highlights it uses a “representative index sampling strategy”^{7,8}. It does not always hold all 500 stocks of the S&P 500.

Our results suggest fundamental risk is not a major concern in our setting. We show there is no positive relation between the incidence of arbitrage opportunities and price deviations from the NAV of either ETF. Moreover, the error correction results suggest that price deviations between the two ETFs also result in convergence. When the price of each ETF deviates from that of the S&P 500 index or the other ETF the error correction mechanism pulls the ETF price back to the index level (other ETF price) in the next minute. Buyer-initiated trades are more common than seller-initiated trades in the underpriced ETF and seller-initiated trades are more prevalent than buyer-initiated trades in the overpriced ETF. This, together with the rapid removal of arbitrage profits (median time of 1-2 minutes), is further evidence that price deviations are seen as worth pursuing by investors.

The ETF providers allude to mispricing in their prospectus. The IVV (2010) prospectus states (p. 3): “The trading prices of the Fund’s shares fluctuate continuously throughout trading

⁶ 2010 SPDR Prospectus, page 45. <https://www.spdrs.com/product/fund.seam?ticker=spy>

⁷ 2010 iShare Prospectus, page 1. http://us.ishares.com/product_info/fund/overview/IVV.htm

⁸ There are also differences in the treatment of dividends received on the underlying stock. SPY retains these as cash until the quarterly distributions. IVV is permitted to reinvest these in S&P 500 stocks until it is scheduled to distribute them. <http://etfdb.com/2010/closer-look-at-sp-500-options/>

hours based on market supply and demand rather than NAV. The trading prices of the Fund's shares may deviate significantly from NAV during periods of market volatility...However, because shares can be created and redeemed ... at NAV... large discounts or premiums to the NAV are not likely to be sustained over the long term.” The SPY (2011) prospectus states (p. 13): “the public trading market price ... may be different from the NAV ... due, in large part, to the fact that supply and demand forces at work in the secondary trading market ... is closely related to, but not identical to, the same forces influencing the price of Index Securities trading individually.” Our results support the notion that there is increased mispricing when the market is more volatile.

The ETF mispricing has attracted a lot of attention in the recent literature. Dolvin (2010) shows the closing prices of the IVV and SPY sometimes deviate enough to allow arbitrage profits. Richie, Daigler, and Gleason (2008) find the SPY can be used to exploit mispricing between the S&P 500 cash index and futures contracts.⁹ They show there is more mispricing when volatility is high. Others, such as Engle and Sarkar (2006), Aber, Li, and Can (2009), and Ackert and Tian (2010) report fluctuations in the mispricing or tracking error of ETFs relative to their underlying index. Elton, Gruber, Comer, and Li (2002) note this mispricing is minimized by the creation and redemption of units via “in kind” transactions. Our contribution is the insight we provide into the microstructure of arbitrage using two close-substitute ETFs.

The mispricing of ETFs has recently caught the attention of regulators. As the joint SEC / CFTC Flash Crash report¹⁰ notes, ETFs accounted for 70% of all US-listed securities that declined by 60% or more during the May 6, 2010 Flash Crash. We hope this paper provides useful insight into the events that precede ETF mispricing.

⁹ Other papers, such as Chu and Hsieh (2002), and Switzer, Varson, and Zghidi (2000) find the introduction of the SPY resulted in less mispricing between the S&P 500 cash and futures indices.

¹⁰ <http://www.cftc.gov/stellent/groups/public/@otherif/documents/ifdocs/opa-jointreport-sec-051810.pdf>

The remainder of this paper is organized as follows: Section 2 contains a description of our data and the rules we use to identify divergent prices. Our results are presented in Section 3 and Section 4 concludes the paper.

2. Data and Arbitrage Identification

We source high-frequency data from the Thompson Reuters Tick History (TRTH) database, which we access via the Securities Institute Research Centre of Asia-Pacific (SIRCA). TRTH data are used by academic researchers (e.g. Fong, Holden, and Trzcinka, 2011), central banks, hedge funds, investment banks, and regulators. More information on TRTH is available at their website.¹¹ We follow Hendershott, Jones, and Menkveld (2010) and limit our analysis to the post-decimalization period of February 2001 onwards. Our sample finishes in August 2010. We use data from the primary listing venue for each ETF. This was the AMEX prior to its takeover by NYSE Euronext in 2008 and NYSE Arca following that point.

It is clear the S&P 500 ETFs diverge from the S&P 500 index on an intraday basis. Figure 1 shows deviations from parity (using one-minute intervals) between the IVV and S&P 500 index during the 1.30pm – 2.30pm period on September 29, 2008. This was a particularly volatile day of trading so the deviations are larger and more frequent than those on some other days, but this example does serve to illustrate that the ETFs frequently move from trading at a premium to a discount versus the S&P 500 index. While the ETF price deviations from the S&P 500 index depicted in Figure 1 are a necessary condition for arbitrage opportunities to exist they are not

¹¹

http://thomsonreuters.com/products_services/financial/financial_products/quantitative_research_trading/tick_history

sufficient. For this there needs to be deviations between ETFs that can be exploited. We outline the algorithm used to test for these arbitrage results below.

[Insert Figure 1 Here]

We adopt the data cleaning approach advocated by Schultz and Shive (2010), which involves deleting observations with any of the following characteristics:

1. Quotes posted outside normal trading hours.
2. Quotes posted in the first and last five minutes of regular trading.¹²
3. Bid price greater than or equal to the corresponding ask price for the same ETF.
4. Ask price four times or more larger than the bid price for the same ETF.
5. Bid-bid or ask-ask return for the same ETF is greater than 25% or less than -25%.
6. Bid price SPY / Ask price IVV > 1.5 or bid price IVV / ask price SPY > 1.5.

We also delete arbitrage observations that appear on May 6, 2010, the day of the Flash Crash, and any observations associated with dividend distributions of ETFs.¹³

Our assumption that trades only take place at the firm quotes in the market is conservative. As Schultz and Shive (2010) note, NYSE trades frequently occur at better prices. There are approximately 553 million SPY quotes and 220 million IVV quotes during the trading hours outlined above that pass our data cleaning process. This represents approximately 36,000 and 14,000 quotes per hour for the SPY and IVV respectively.

¹² This is broadly consistent with Schultz and Shive (2010) who omit quotes from the first and last four minutes of trading.

¹³ ETFs typically distribute dividends, which they have collected from the stocks they hold, on a quarterly basis. These distributions rarely occur on the same day, which means that illusionary profit opportunities appear to be available in the period until both ETFs have gone ex-distribution.

Our profit results are or are close to “net profits.” The spread is accounted for and market impact costs are minimal for most traders with such highly liquid instruments. The only remaining execution costs appear to be commissions and short-selling costs, which are typically ignored in the pairs trading literature. Goldstein, Irvine, Kandel, and Wiener (2009) show the average one-way commission has declined over time to less than five cents per share in 2004. These have declined further since this point with 2010 NYSE documentation stating commissions of 0.15 – 0.3 cents per share for those with direct access to NYSE Arca.¹⁴ Moreover, D’Avolio (2002) estimate S&P 500 stocks can be borrowed at less than 0.25% per annum, or approximately 0.0007% per day. Given these minimal costs, we suggest the profits we document are close to the actual profits large hedge funds / institutions could earn by exploiting these arbitrage opportunities. However, we apply a profit threshold of 0.2% as a conservative measure to ensure our sample is not dominated by the numerous smaller divergences.

The speed at which orders can be executed has increased over time. Bacidore, Ross, and Sofianos (2003) find the average order-to-execution time for NYSE orders in 1999 was 22.5 seconds. Garvey and Wu (2009) document a mean (median) execution time of 12 (4) seconds for the 1999-2003 period. Hendershott and Moulton (2009) show execution times declined to less than one second in 2006/2007 and Hasbrouck and Saar (2010) find some traders now react to market events within 2-3 milliseconds. However, we take the conservative approach and assume the orders take 15 seconds to be executed. For quotes to be included as valid quotes, both ETF quotes need to have been updated in the last five minutes. Our algorithm is as follows:

1. If $\text{bid SPY} / \text{ask IVV} \geq 1.002$ or $\text{bid IVV} / \text{ask SPY} \geq 1.002$ a *potential* arbitrage opening trade is identified.

¹⁴ http://www.nyse.com/pdfs/NYSE_Arca_Marketplace_Fees_11_5_2010-Clean.pdf

2. The *actual* arbitrage opening trade is opened at the first set of quotes for each ETF that appear 15 seconds after the *potential* opening arbitrage trade is identified. We assume conditional limit, “fill or kill” orders are used that equate to a minimum opening profit of 0.2%.
3. In the case of the short SPY / long IVV trade, SPY ask and IVV bid prices are monitored. When a situation of $\text{IVV bid} / \text{SPY ask} \geq 1$ is observed a *potential* arbitrage closing trade is identified. If a short IVV / long SPY trade was opened IVV ask and SPY bid prices are monitored and a *potential* arbitrage closing trade is identified when $\text{SPY bid} / \text{IVV ask} \geq 1$. If 0.2% profit has been secured when the trade is opened we assume an arbitrageur is happy to close the trade at prices that do not result in further profit.
4. The *actual* arbitrage closing trade occurs at the first set of quotes for each ETF that appear 15 seconds after the *potential* closing arbitrage trade is identified. We again assume conditional limit, “fill or kill” orders are used that equate to a minimum closing profit of 0.2%.

We also test scenarios where an arbitrageur requires an opening profit of 0.3% and 0.4% and is happy taking a closing loss of up to 0.1% and 0.2% respectively to leave overall profits of 0.2% or more. These two strategies require less than total price convergence for arbitrages to be closed so are likely to be able to be closed more quickly.

Examples of arbitrage opportunities included in our results are provided in Figure 2a, b, and c. As Figure 2a shows, at 14.40 on November 11, 2002, the bid price of the SPY was \$84.08 and the ask price of the IVV was \$83.82. These prices prevailed for 15 seconds so we assume an arbitrageur could sell at the SPY bid and buy at the IVV ask and lock in profit of 0.31% $(84.08 - 83.82) / 0.5 \times (84.08 + 83.82)$, which is realized at convergence. The arbitrageur goes long the IVV and shorts the SPY. She maintains this position until 15.06 when the IVV bid and SPY ask

converge at 83.50 for a period longer than 15 seconds. At this point, positions are closed at no further profit or loss, leaving overall profits of 0.31%. It is important to note this percentage profit calculation (and all others in this paper) is based on the assumption of Mitchell, Pulvino, and Stafford (2002), which is that a 50% margin (which earns no return) is required for both long and short positions. Actual profits would be larger if a lower margin is required and / or a return is earned on the margin that is posted. In 2002 when the arbitrage opportunity depicted in Figure 2a was created the spread on the IVV was relatively large compared to that of the SPY. At 14.40 when the arbitrage was opened the IVV (SPY) spread was 0.24% (0.01%). The IVV spread declined over time, which is shown in Figures 2b and c.

[Insert Figure 2a, b, and c Here]

3. Results

Table 1 contains correlations, error correction model results, and liquidity information for the SPY and IVV. Panel A shows that both movements in both ETFs are highly correlated with those of the S&P 500. The SPY – S&P 500 daily correlation is 0.98 while the IVV – S&P 500 correlation is 0.99. Engle and Granger (1987) state an error-correction representation exists if a set of variables are co-integrated. The unreported Johansen (1991) co-integration test shows that ETF prices and S&P500 at both the daily and one minute frequencies are co-integrated. We estimate the error correction model at both the daily and one minute frequencies as follows in equations 1a-1c.

$$\Delta SPY_t = \alpha + \beta \Delta S\&P500_{t-1} + \gamma (SPY_{t-1} - \delta S\&P500_{t-1}) + \varepsilon_t \quad (1a)$$

$$\Delta IVV_t = \alpha + \beta \Delta S\&P500_{t-1} + \gamma (IVV_{t-1} - \delta S\&P500_{t-1}) + \varepsilon_t \quad (1b)$$

$$\Delta SPY_t = \alpha + \beta \Delta IVV_{t-1} + \gamma (SPY_{t-1} - \delta IVV_{t-1}) + \varepsilon_t \quad (1c)$$

Where $\gamma(SPY_{t-1} - \delta S\&P500_{t-1})$, $\gamma(IVV_{t-1} - \delta S\&P500_{t-1})$ and $\gamma(SPY_{t-1} - \delta IVV_{t-1})$ are the error correction terms that describes how the ETFs prices and S&P500 vary in short-run are consistent with a long-run co-integrating relationship; δ s are co-integration parameters. The statistical significance of γ indicates that both ETFs error correct back to the S&P 500 at each of these frequencies.

The Panel B results show the daily SPY – IVV correlation is 0.99. Both ETFs error correct back to each other at the daily frequency and one minute frequency, although the error correction is much stronger in the IVV. Panel C quantifies just how liquid these ETFs are. During our sample period the SPY had an average daily value traded of \$14,711 million, while the IVV had an average daily value traded of \$220 million. The SPY was more liquid than any NYSE / Amex / Nasdaq stock and the IVV was in the 99th percentile of stocks. The average daily quoted spread of the SPY (IVV) is just 0.016% (0.045%), and the average daily effective spread of the SPY (IVV) is just 0.031% (0.12%).

[Insert Table 1 Here]

Profit results are presented in Table 2. From Table 2 Panels A and B it is clear that it is almost equally likely for each ETF to be underpriced (overpriced). The SPY (IVV) is underpriced (overpriced) on 90 occasions and the IVV (SPY) is underpriced (overpriced) 93 times. The mean

profits are also very similar (0.33% versus 0.32%). The maximum profit is 1.87% when the SPY (IVV) is underpriced (overpriced) and 1.27% when the IVV (SPY) is underpriced (overpriced).

We also test two variations of the base strategy and report these results in Appendix 1. These variation strategies require profits of 0.3% and 0.4% respectively for trades to be opened and convergence equating to losses of no more than -0.1% and -0.2% respectively for trades to be closed. All three strategies therefore identify arbitrage opportunities with profits of 0.2% and larger. The results for these strategies are not mutually exclusive. For instance, an arbitrage opportunity has a profit 0.5% will appear in the base scenario and each of the variation strategies.¹⁵ The results suggest profits are higher, on average, when the 0.3% and 0.4% opening thresholds are applied, but this higher limit results in fewer opportunities. For example, the mean profits of the 0.3% / -0.1% or 0.4% / -0.2% strategies when the SPY (IVV) is underpriced (overpriced) are 0.36% and 0.42% compared to 0.33% for the base strategy.

Panels C and D contain the average profit across all arbitrage opportunities by minute. Minute 0 refers to the profit on offer at the start of the minute when the arbitrage occurs. The means of 0.1% and 0.08% reflect the fact that the divergence required to open the arbitrage (0.2%) has not occurred yet. This occurs at various stages during minute 0. Minute 1 (-1) is the profit on offer at the start of the minute following (preceeding) the minute where the 0.2% arbitrage threshold is reached. Statistically significant average arbitrage profits exist at the start of minute 1 but these have disappeared by the start of minute 2. The negative profits prior to time 0, which are statistically significant at times, do not indicate that positive profits could have been made from an opposite strategy of buying (selling) the overpriced (underpriced) ETF. They

¹⁵ Each arbitrage is opened and closed within normal trading hours (excluding the first and last 5 minutes of trading) at quotes that prevail 15 seconds after each of the criteria are met. Our results are conservative as we only consider arbitrages generated with quotes posted in normal trading hours (excluding the first and last five minutes). In reality arbitrageurs could exploit divergent quotes after-hours and improve their profitability.

simply indicate that, on average, at that point in time the ask price of the underpriced ETF was higher than the bid price of the overpriced ETF.

We present summary statistics for the length of time that divergent quotes allowing arbitrage profits persist for in Table 2 Panels A and B. Each duration is measured as the divergent pricing lasts for plus 15 seconds (the delay that is imposed once the arbitrage opportunity is first identified). This ensures it is a conservative indication of the length of time an arbitrageur has to actually exploit the opportunity.

The divergent pricing which creates the arbitrage opportunities is removed relatively quickly. The median duration is 2.27 minutes when the SPY (IVV) is underpriced (overpriced) and 0.92 minutes when the IVV (SPY) is underpriced (overpriced). Arbitrages involving SPY (IVV) under pricing (overpricing) tend to last longer. The 75th percentile duration is 21.25 minutes when SPY is underpriced compared to just 1.64 minutes when IVV is underpriced. There are some extreme observations, due to arbitrage opportunities occurring at the end of the day (especially on Friday) which inflates the mean duration of profit opportunities. For instance, the maximum duration of 3,679 minutes was opened in the afternoon of Friday February 3, 2006 and closed on the morning of Monday February 6, 2006. However, even with these included the overall mean duration for the base strategy is 4 – 86 minutes.

To our knowledge, the rapid removal of arbitrage opportunities is quite different to the results of other studies on arbitrage in equity markets. For example, Froot and Dabora (1999) show mispricing in stocks listed in different locations can prevail for over four years¹⁶, Mitchell, Pulvino, and Stafford (2002) show divergence in the price of a parent company and listed subsidiary can last for over five months¹⁷, and Schultz and Shive (2010) show mispricing in class

¹⁶ See Froot and Dabora (1999) Figure 1 on page 193.

¹⁷ See Mitchell, Pulvino, and Stafford (2002) Figure 1 on page 568.

of stock with different voting rights can persist for two years.¹⁸ However, the relatively quick restoration of efficient prices is consistent with Busse and Green (2002). They show prices converge to efficient levels following CNBC reports in one to fifteen minutes depending on whether the report is good or bad news. Moreover, Chordia, Roll, and Subrahmanyam (2005) find investors take between five and sixty minutes to restore prices to efficient levels following order imbalances.

[Insert Table 2 Here]

In Table 3 we present results which relate to the relation between changes in the NAVs of the ETFs, and arbitrage opportunities and changes in levels of the VIX and arbitrage opportunities. The IVV (2010) prospectus (p. 3) states: “The trading prices of the Fund’s shares fluctuate continuously throughout trading hours based on market supply and demand rather than NAV. The trading prices of the Fund’s shares may deviate significantly from NAV during periods of market volatility...” In Panel A we present results for a regression of the absolute daily NAV deviations of the SPY and IVV on a dummy variable that equals 1 if there is an arbitrage opportunity that day and 0 otherwise. The coefficient of this variable is negative which suggests that arbitrage opportunities occur, on average, on days when NAV deviations are lower than average. Changes in NAV are not the driver of the arbitrage opportunities we document. The Panel B results relate to a regression of VIX on a dummy variable that equals 1 on days when arbitrage opportunities occur and 0 otherwise. This shows that the level of the VIX is 6 percentage points higher than its 22.5 percent average on days when arbitrages occur. This result

¹⁸ See Schultz and Shive (2009) Figure 2 on page 4.

confirms that it is the temporary deviation of prices from NAV during more volatile market conditions that result in ETF mispricing.

[Insert Table 3 Here]

We now turn our attention to considering market characteristics before each arbitrage opportunity is created. We calculate each variable at the time of the event. Time $t-1$ is the start of the minute prior to the arbitrage opportunity, $t-2$ is the previous minute, and so on. We measure each variable of the event day and on the previous 20 trading days. The percentage change between the variable level on the event day and the mean on the previous 20 days is then calculated.¹⁹ The median percentage change across all events is presented in Table 4. We have a pre-event period of $t-5$ to $t-2$, an event period of $t-1$ to $t+1$, and a post-event period of $t+2$ to $t+5$. Variables that are statistically significantly different to the previous period at the 10% level or better are in bold. For instance, if the $t-1$ to $t+1$ variable is different to the $t-5$ to $t-2$ variable the $t-1$ to $t+1$ number is in bold.

Spread is calculated for the first trade in each interval as two times the absolute value of the difference between the transaction price and the prevailing mid-price (e.g. Goyenko, Holden, and Trzcinka, 2009). Depth is the value of shares at the first level of the book (both the bid and ask) size at the start of each interval. Order Imbalance is calculated as the difference in the absolute value between buyer-initiated trades and seller-initiated trades divided by the sum of the two in each one-minute interval. Trades are classified as buyer- or seller-initiated using the Lee and Ready (1991) algorithm. As per Sarkar and Schwartz (2009), sidedness is the correlation between the standardized number of buyer-initiated trades and seller-initiated trades in the last 30

¹⁹ Change is calculated for the sidedness as this is already expressed as a percentage.

minutes leading up to the arbitrages. The Return Standard Deviation is calculated based on the standard deviation of traded returns in each one-minute interval. Trade Value is the total value of trades during each minute interval, and Spread Standard Deviation is computed as the standard deviation of effective spreads in each interval.

The Table 4 results show that arbitrage opportunities occur at times when liquidity declines in both the SPY and IVV. Spreads are 41% - 51% higher in the event window (minute -1 to +1) than the equivalent time of day in the prior 20 trading days in the SPY. When the IVV (SPY) is underpriced (overpriced) spreads are 16%-18% larger. Moreover, three of the four spreads are statistically significantly larger in the $[-1, +1]$ interval than they are in the $[-5, -2]$ interval. Depth is statistically significantly lower in the event window than $[-5, -2]$ interval and is 18% lower than the level from the previous 20 days in the IVV when the IVV is underpriced, but this decline in depth is not evident in other instances.

Order imbalance increases dramatically in the minutes surrounding an arbitrage opportunity. It is 17% - 41% higher in the $[-1, +1]$ interval on event days than the previous 20 trading days and the event window order imbalance is higher than that of the previous interval in three out of the four instances. This is consistent with markets becoming more one-sided. The sidedness measure of Sarkar and Schwartz (2009) also indicates markets are more one-sided at the point at which arbitrage opportunities are created. However, there is no indication of stronger sidedness (based on the Sarkar and Schwartz (2009) measure) in the interval $[-1, +1]$.

There is a surge in spread standard deviation in the IVV in the minutes surrounding arbitrage opportunities. These are 280% (300%) higher in the event window than they are at the same time in prior 20 trading days when the IVV is overpriced (underpriced). The spread standard deviation of the SPY also increases (24% - 61%) when arbitrage opportunities are created. Trading value in both the SPY and IVV also increases around arbitrage opportunities.

Johnson (2008) points out that numerous empirical studies find volume and liquidity (e.g. spread) to be unrelated. He develops a frictionless model to explain this phenomenon. His model shows that liquidity reflects the risk bearing capacity or average willingness of traders to accommodate trades while volume represents the second moment of liquidity changes or liquidity risk. Pástor and Stambaugh (2003) and Acharya and Pederson (2004) highlight the importance of liquidity risk on asset pricing. Johnson (2008, p. 389) suggests that “liquidity risk is important from a policy perspective because of the danger posed by large drops in liquidity, which may lead to price distortions, disruptions in risk transfer, and possibly inefficient liquidation of real investments.” Our results indicate the significant price distortion between the two ETFs occurs when liquidity dries up and liquidity risk increases, which is a period that the risk bearing capacity drops and such capacity is likely to fall further. We also find that the return standard deviation increases.

The results for the interval $t+2$ to $t+5$ indicate that spreads stay relative high in the minutes following each arbitrage opportunity, order imbalance, and return standard deviation declines in the SPY but not the IVV, and the spread standard deviation declines in the IVV but not the SPY.

[Insert Table 4 Here]

Divergent pricing in ETFs has recently captured the attention of regulators. The May 6, 2010 “Flash Crash” has been the catalyst for regulators to look at the reasons behind ETF pricing deviations from the indices they seek to mirror. In their joint report²⁰, the SEC and CFTC note

²⁰ The report is entitled “Preliminary Findings Regarding the Market Events of May 6, 2010” and is available here: <http://www.cftc.gov/stellent/groups/public/@otherif/documents/ifdocs/opa-jointreport-sec-051810.pdf>

ETFs made up 70% of all US-listed securities that plunged 60% or more on May 6. The SEC / CFTC report suggests one explanation for this might be institutions using ETFs as a quick way of reducing market exposure. Another possible explanation is the use of stop-loss market orders in ETFs by individual investors. Borkovec, Domowitz, Serbin, and Yegerman (2010, p. 1) suggest the failure in ETF price discovery during this event was due to “an extreme deterioration in liquidity, both in absolute terms and relative to individual securities in the baskets tracked by the funds.”

We plot the deviations of the SPY and IVV from the level of the S&P 500 index at one-minute intervals during the 2.30pm – 3.30pm period on May 6, 2010 in Figure 3. Numbers greater (less) than zero indicate the ETF is trading at a premium (discount). The black bars represent the total SPY variation and the entire white bar (including the black portion) represents the IVV variation. Figure 3 makes it clear the IVV was more affected than the SPY. Its deviations from parity range from +4% to -10%. It is also evident the ETFs change from trading at a large premium one minute (2.45pm) to a large discount the next minute (2.46pm). On one occasion (2.49pm) the SPY is trading at a premium of over 2% and the IVV is trading at a discount of more than -2%. We do not include any observations from May 6, 2010 in our formal analysis.

[Insert Figure 3 Here]

In Table 5 we present results which provide insight into the convergence of prices following the arbitrage opportunities. We consider whether there is a difference in the number, volume, and dollar volume of buyer- and seller-initiated trades in the time between mispricing allowing arbitrage profits and their removal. We also look at aggressive buy (sell) trades, which we define as trades that go through at or above the prevailing ask price (at or below the prevailing

bid price). The results indicate buyer-initiated trades are more likely in the underpriced ETF and seller-initiated trades are more common in the overpriced ETF. Moreover, aggressive buy trades are more prevalent in the underpriced ETF and aggressive sell trades are more likely in the overpriced ETF. The differences are statistically significant when the IVV is underpriced and the SPY is overpriced. Taken together, we suggest the Table 5 results provide evidence that arbitrageurs act to profit from the mispricing. Their actions lead to an ETF prices error correcting back to parity and short-lived mispricing.

[Insert Table 5 Here]

4. Conclusions

We document the intraday characteristics that prevail when mispricing which is large enough to create arbitrage profits. We show that liquidity declines (i.e. spreads increase) and order imbalance increases. The standard deviation of spread changes increase which suggests that liquidity risk increases. Trade value increases which, according to Johnson (2008), is also indicative of an increase in liquidity volatility. Return volatility is also higher when arbitrage opportunities occur.

Our results are based on two S&P 500 ETFs. The SPDR Trust (ticker SPY) is more liquid than any US stock and the iShare (ticker IVV) is in the 99th percentile of stocks based on liquidity. Having highly liquid instruments is important as it allows us to calculate meaningful statistics at high frequencies. The S&P 500 ETFs are natural instruments for other reasons. They are close substitutes for each other (low fundamental risk). The ETF providers have an incentive to minimize tracking error so as to attract more funds and holders can exchange the ETF for the

underlying shares, which mitigates the risk that these ETFs will not converge following mispricing. We find that the ETFs have a daily return correlation of 0.99 and their deviations correct back following mispricing. Median durations are just 1-2 minutes. There is an increase in buyer-initiated trades (seller-initiated trades) in the underpriced (overpriced ETF) when mispricing occurs as arbitragers seek to profit from the mispricing. NAV differences between the ETFs do not drive the mispricing. Arbitrage opportunities are, however, more likely to be created when the market is more volatile.

The pricing of ETFs have recently captured the attention of regulators. During the May 6, 2010 Flash Crash a disproportionate number of ETFs were among those securities that experienced the largest price declines. The prices of ETFs also temporarily diverged by large amounts (greater than 10%) from their “true value” based on the indices they track. We hope that our research improves understanding of the factors that result in ETF mispricing on an intraday basis.

References

- Aber, Jack W., Li, Dan., and Luc Can. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management*, 10, 210-221.
- Abreu, Dilip and Markus K. Brunnermeier. (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics*, 66, 341-360
- Acharya, Viral .V. and Lasse Heje Pedersen. (2004). Liquidity risk and asset pricing. *Journal of Financial Economics*, 77, 375-410.
- Ackert, Lucy F. and Yisong S. Tian. (2000). Arbitrage and valuation in the market for Standard and Poor's Depositary Receipts. *Financial Management*, 29(3), 71-87.
- Bacidore, Jeffrey, Ross, Katharine., and George Sofianos. (2003). Quantifying market order execution quality at the NYSE. *Journal of Financial Markets*, 6, 281-307.
- Bondarenko, Oleg. (2003). Statistical arbitrage and securities prices. *Review of Financial Studies*, 16(3), 875-919.
- Borkovec, Milan., Domowitz, Ian., Serbin, Vitaly., and Henry Yegerman. (2010). Liquidity and price discovery in exchange traded funds: One of several possible lessons from the flash crash. ITG Investment Group. http://www.itg.com/news_events/ITG-Paper-LiquidityPriceDiscovery.pdf
- Busse, Jeffrey and Clifford T. Green. (2002). Market efficiency in real time. *Journal of Financial Economics*, 65, 415-437.
- Chu, Quentin C and Wen-Liang Gideon Hsieh. (2002). Pricing Efficiency of the S&P 500 index market: Evidence from the Standard and Poor's Depositary Receipts. *Journal of Futures Markets*, 22(9). 877-900.
- D'Avolio, Gene. (2002). The market for borrowing stock. *Journal of Financial Economics*, 66 (2-3), 271-306.
- De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J Waldmann. (1990). Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-738.
- Deuskar, Prachi and Timothy Johnson (2011). Market liquidity and flow driven risk. *Review of Financial Studies* 24(3), 721-753.
- Dolvin, Steven D. (2010). S&P ETFs: Arbitrage opportunities and market forecasting. *The Journal of Index Investing*, Summer, 107-116.
- Elton, Edwin J., Gruber, Martin J., Comer, George, Kai Li. (2002). Spiders: Where are the bugs? *Journal of Business*, 75(3), 453-472.

- Engle, Robert, F. and Clive W. J. Granger (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55, 251-276.
- Engle, Robert and Debojyoti Sarkar. (2006). Premiums-discounts and exchange traded funds. *Journal of Derivatives*, Summer, 27-45.
- Engleberg, Joseph, Gao, Pengjie, and Ravi Jagannathan. (2010). *An anatomy of pairs trading: the role of idiosyncratic news common information and liquidity*. SSRN Working Paper: <http://ssrn.com/abstract=1330689>
- Fong, Kingsley, Craig Holden, and Charles Trzcinka (2011). What are the best liquidity proxies for global research? SSRN working paper, <http://ssrn.com/abstract=1558447>.
- Froot, Kenneth. A. and Emil. M. Dabora. (1999). How are stock prices affected by the location of trade? *Journal of Financial Economics*, 53, 189-216.
- Garvey, Ryan and Fei Wu (2009) Intraday time and order execution quality dimensions. *Journal of Financial Markets*, 12, 203-228.
- Gagnon, Louis and G. Andrew Karolyi. (2010). Multi-market trading and arbitrage. *Journal of Financial Economics*, 97, 53-80.
- Gatev, Evan, Goetzmann, William N. and K. Geert Rouwenhorst. (2003). Pairs trading: Performance of a relative value arbitrage rule. *Review of Financial Studies*, 19(3), 797-827.
- Goldstein, Michael A., Irvine, Paul., Kandel, Eugene., and Zvi Wiener. (2009). Brokerage commissions and institutional trading patterns. *Review of Financial Studies*, 22(12), 5175-5212.
- Hamm, Sophia J.W., (2011). The effect of ETFs on stock liquidity. SSRN Working Paper: <http://ssrn.com/abstract=1687914>
- Hasbrouck, Joel and Gideon Saar. (2010). Low-latency trading. SSRN Working Paper: <http://ssrn.com/abstract=1695460>
- Hendershott, Terrence., Jones, Charles M., and Albert J. Menkveld. (2010). Does algorithmic trading improve liquidity? *Journal of Finance – Forthcoming*.
- Hendershott, Terrence and Pamela C. Moulton. (2009). *Speed and stock market quality: The NYSE's hybrid*. SSRN Working Paper: <http://ssrn.com/abstract=1159773>
- Hogan, Steve., Jarrow, Robert., Teo, Melvyn., and Mitch Warachka. (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, 73, 525-565.
- IVV (2010). 2010 Prospectus to Shareholders. iShares S&P 500 Index Fund, August 2010.

- Johansen, Søren. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551–1580.
- Johnson, Timothy C. (2008). Volume, liquidity, and liquidity risk. *Journal of Financial Economics*, 87, 388-417.
- Lee, Charles, and Mark Ready. (1991). Inferring trade direction from intraday data, *Journal of Finance*, 46, 733–746.
- Mitchell, Mark., Pulvino, Todd., and Erik Stafford. (2002). Limited arbitrage in equity markets. *Journal of Finance*, 57(2), 551-584.
- Pástor, Ľuboš, and Robert F. Stambaugh (2003). Liquidity risk and expected stock returns, *Journal of Political Economy*, 111, 642–685.
- Richie, Nivine, Daigler, Robert T., and Kimberly C. Gleason. (2008). The limits to stock index arbitrage: Examining S&P 500 Futures and SPDRs. *Journal of Futures Markets*, 28(12), 1182-1205.
- Sarkar, Asani, and Robert A. Schwartz (2009). Market sidedness: Insights into motives for trade initiation. *Journal of Finance*, 64(1), 375-423.
- Schultz, Paul, and Sophie Shive. (2010). Mispricing of dual-class shares: Profit opportunities, arbitrage, and trading. *Journal of Financial Economics*, 98, 524-549.
- Shleifer, Andrei. and Robert W. Vishny. (1997). The limits of arbitrage. *Journal of Finance*, 52, 35-55.
- SPY (2011). SPDR S&P 500 ETF Trust Prospectus, January 2011.
- Switzer, Lorne N., Varson, Paula L., and Samia Zghidi. (2000). Standard and Poor's Depositary Receipts and the performance of the S&P 500 index futures market. *Journal of Futures Markets*, 20(8), 705-716.
- Xiong, Wei. (2001). Convergence trading with wealth effects: An amplification mechanism in financial markets. *Journal of Financial Economics*, 62, 247–292

Table 1
Correlation, Error Correction Results, and Liquidity

<i>Panel A: ETF and S&P 500 Movements</i>					
	Correlation	Daily Error Correction		1-Minute Error Correction	
		Coefficient	<i>t</i> -statistic	Coefficient	<i>t</i> -statistic
SPY	0.98	-0.01	-1.93	-0.05	-134.97
IVV	0.99	-0.47	-3.65	-0.18	-395.27
<i>Panel B: ETFs Movements</i>					
	Correlation	Daily Error Correction		1-Minute Error Correction	
		Coefficient	<i>t</i> -statistic	Coefficient	<i>t</i> -statistic
SPY	0.99	-0.05	-2.17	-0.010	-39.52
IVV		-0.04	-1.88	-0.199	-419.83
<i>Panel C: ETF Liquidity</i>					
	Value	Value Percentile	No. Trades	Quoted Spread	Effective Spread
SPY	14,711	Max	203,974	0.016%	0.031%
IVV	220	99	5,648	0.045%	0.120%

This table contains correlation, error correction, and liquidity results for the February 2001 – August 2010 period. Daily data in Panels A and B are from CRSP. Intraday data in these two panels are from the Thompson Reuters Tick History (TRTH) database. Panel C is based on daily CRSP data.

Table 2
Arbitrage Profits and Durations

<i>Panel A: SPY Underpriced / IVV Overpriced - Profit and Duration Summary Statistics</i>								
	N	Min	25 Per	Median	Mean	75 Per	Max	Std Dev
Profit	90	0.20	0.23	0.27	0.33	0.34	1.87	0.22
Duration	90	0.32	0.76	2.27	86.46	21.25	3,769.24	404.95
<i>Panel B: IVV Underpriced / SPY Overpriced - Profit and Duration Summary Statistics</i>								
	N	Min	25 Per	Median	Mean	75 Per	Max	Std Dev
Profit	93	0.20	0.24	0.27	0.32	0.33	1.27	0.15
Duration	93	0.06	0.52	0.92	3.92	1.64	186.96	19.57
<i>Panel C: SPY Underpriced / IVV Overpriced - Mean Profit By Minute</i>								
Minute	-5	-4	-3	-2	-1	0	1	2
Mean	0.00	-0.02	-0.01	0.01	0.01	0.10	0.09	0.03
<i>t</i> -statistic	-0.34	-1.10	-0.51	0.45	0.95	6.64	4.10	1.25
<i>Panel D: IVV Underpriced / SPY Overpriced - Mean Profit By Minute</i>								
Minute	-5	-4	-3	-2	-1	0	1	2
Mean	-0.06	-0.07	-0.05	-0.04	-0.01	0.08	0.06	0.00
<i>t</i> -statistic	-5.71	-5.99	-4.69	-3.86	-1.15	5.65	3.79	0.52

The data are from the Thompson Reuters Tick History (TRTH) database and the results relate to the February 2001 – August 2010 period. All profits are in percent and durations are in minutes. *t*-statistics that are statistically significant at the 10% level or higher are in bold.

Table 3
Arbitrage Opportunities, NAV Deviations, and Market Volatility

Variable	Coefficient	<i>t</i> -Statistic	Adjusted R ²
<i>Panel A: NAV Deviations</i>			
Constant	0.0011	19.13	0.0156
Arbitrage Day	-0.0006	-7.22	
<i>Panel B: VIX</i>			
Constant	22.5040	35.32	0.0178
Arbitrage Day	6.3215	4.08	

The data are from the Thompson Reuters Tick History (TRTH) database and the results relate to the February 2001 – August 2010 period. The Panel A results are based on a regression of the absolute daily NAV deviations on dummy variable that equals 1 if there is an arbitrage opportunity that day and 0 otherwise. The NAV deviation is the end of day NAV difference between SPY and IVV. The Panel B results are based on a regression of VIX dummy variable that equals 1 if there is an arbitrage opportunity that day and 0 otherwise.

Table 4
Market Characteristics

<i>Panel A: SPY Underpriced / IVV Overpriced</i>						
	SPY			IVV		
	<i>t</i> -5 to <i>t</i> -2	<i>t</i> -1 to <i>t</i> +1	<i>t</i> +2 to <i>t</i> +5	<i>t</i> -5 to <i>t</i> -2	<i>t</i> -1 to <i>t</i> +1	<i>t</i> +2 to <i>t</i> +5
Spread	0.103	0.413	0.294	0.257	0.508	0.478
Depth	-0.074	-0.073	-0.042	-0.032	-0.006	-0.038
Order Imbalance	-0.073	0.406	-0.061	0.137	0.167	0.157
Sidedness	-0.056	-0.061	-0.038	-0.040	-0.047	-0.053
Return Std. Dev.	0.179	1.317	0.704	0.884	1.285	1.390
Trade Value	0.337	2.305	1.805	0.460	0.929	0.752
Spread Std. Dev.	0.125	0.612	0.625	1.686	2.810	1.903
<i>Panel A: IVV Underpriced / SPY Overpriced</i>						
	SPY			IVV		
	<i>t</i> -5 to <i>t</i> -2	<i>t</i> -1 to <i>t</i> +1	<i>t</i> +2 to <i>t</i> +5	<i>t</i> -5 to <i>t</i> -2	<i>t</i> -1 to <i>t</i> +1	<i>t</i> +2 to <i>t</i> +5
Spread	0.030	0.176	0.160	0.101	0.159	0.195
Depth	-0.046	-0.045	-0.126	-0.103	-0.179	-0.067
Order Imbalance	0.000	0.271	-0.080	0.154	0.172	0.169
Sidedness	-0.086	-0.108	-0.104	-0.050	-0.042	-0.049
Return Std. Dev.	0.067	0.888	0.446	1.786	1.827	1.375
Trade Value	0.076	0.986	1.120	0.708	1.444	0.559
Spread Std. Dev.	-0.012	0.235	0.184	1.090	3.000	2.080

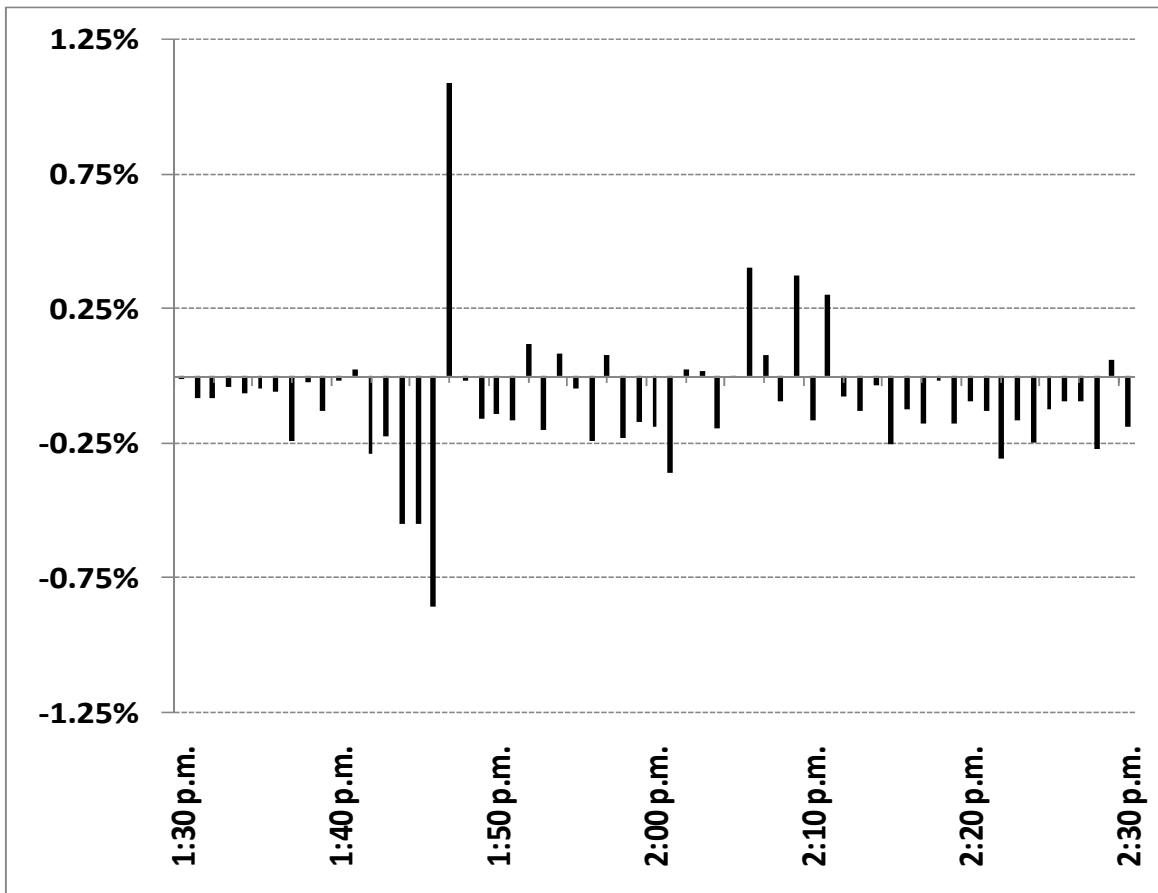
The data are from the Thompson Reuters Tick History (TRTH) database and the results relate to the February 2001 – August 2010 period. . Spread, depth, order imbalance, return standard deviation, trade value, and spread standard deviation are percentage increases or decreases based on the numbers from the same time of the day on the previous 20 trading days. Sidedness is the change from the same time of day on the previous 20 trading days. The statistical significance of the difference of the variables across adjacent windows is calculated using a bootstrap test. Statistical significance at the 10% level or higher is in bold on the column to the right. Therefore bold in column *t*-1 to *t*+1 means these numbers are statistically significantly different to those in *t*-5 to *t*-2. Bold in column *t*+2 to *t*+5 means these numbers are statistically significantly different to those in *t*-1 to *t*+1.

Table 5
Convergence

		<u>SPY Underpriced / IVV Overpriced</u>			<u>IVV Underpriced / SPY Overpriced</u>		
		Number	Volume	\$ Volume	Number	Volume	\$ Volume
SPY	Buy	55%	54%	54%	44%	42%	42%
	Sell	45%	46%	46%	56%	58%	58%
IVV	Buy	50%	48%	48%	63%	58%	58%
	Sell	50%	52%	52%	37%	42%	42%
SPY	Aggressive Buy	50%	51%	51%	34%	36%	36%
	Aggressive Sell	41%	43%	43%	46%	52%	52%
IVV	Aggressive Buy	31%	32%	32%	36%	38%	38%
	Aggressive Sell	40%	41%	41%	22%	25%	25%

The data are from the Thompson Reuters Tick History (TRTH) database and the results relate to the February 2001 – August 2010 period. . Trades are classified based on the Lee and Ready (1991) algorithm. The proportion of buyer- and seller-initiated trades in the time between mispricing allowing arbitrage profits and their removal are presented. Aggressive buy (sell) trades are those trades at or above the prevailing ask price (at or below the prevailing bid price). Statistically significant differences between buy and sell and between aggressive buy and aggressive sell trades (at the 10% level are in bold).

Figure 1
IVV Deviations from the S&P 500, September 29, 2008



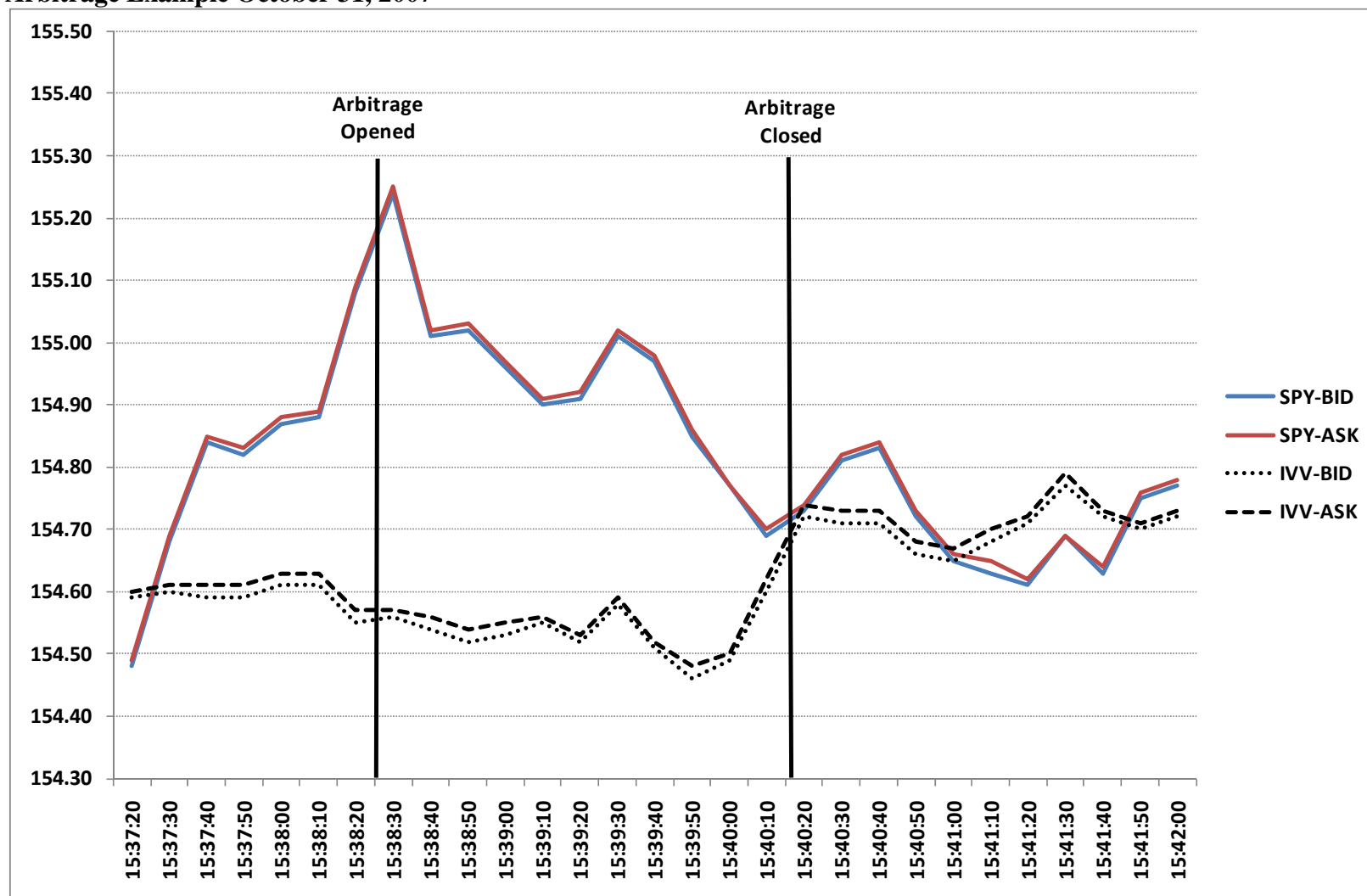
Deviations from parity between the IVV and S&P 500 during one-minute intervals on 29 September 2008. Numbers greater (less) than zero indicate the IVV is trading at a premium (discount).

Figure 2a
Arbitrage Example November 11, 2002



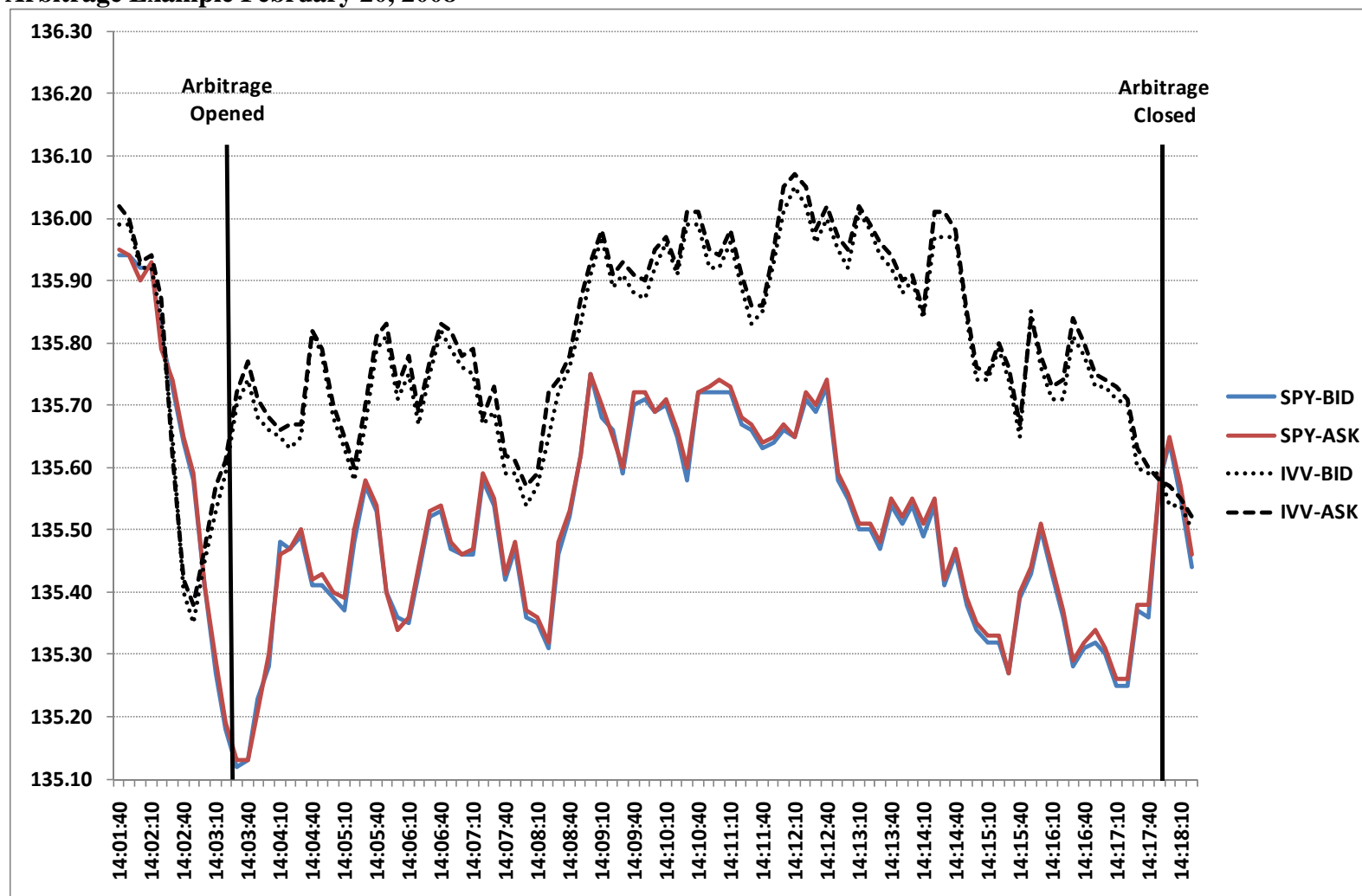
An arbitrage that was opened by selling (buying) the SPY (IVV) at the bid (ask) and closed at by selling (buying) the IVV (SPY) at the bid (ask).

Figure 2b
Arbitrage Example October 31, 2007



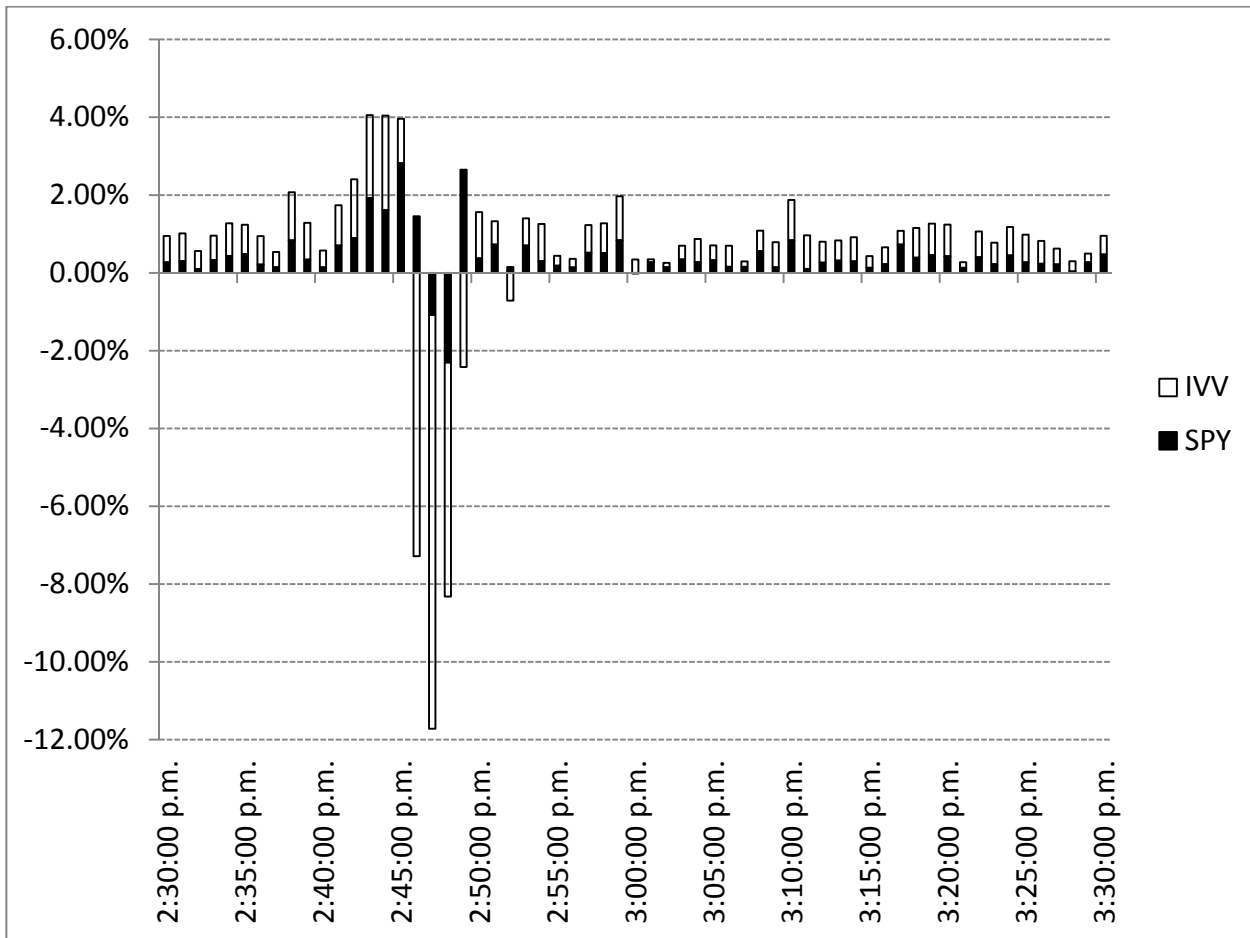
An arbitrage that was opened by selling (buying) the SPY (IVV) at the bid (ask) and closed at by selling (buying) the IVV (SPY) at the bid (ask).

Figure 2c
Arbitrage Example February 20, 2008



An arbitrage that was opened by selling (buying) the IVV (SPY) at the bid (ask) and closed at by selling (buying) the SPY (IVV) at the bid (ask).

Figure 3
IVV and SPY Deviations from the S&P 500 During Flash Crash (May 6, 2010)



Deviations from parity between the IVV and S&P 500 and SPY and S&P 500 during one-minute intervals during the Flash Crash on May 6, 2010. Numbers greater (less) than zero indicate the IVV is trading at a premium (discount).

Appendix 1
Arbitrage Profits for Alternative Strategies

Threshold %	N	Min	Median	Mean	Max	Std Dev
<i>Panel A: SPY Underpriced / IVV Overpriced</i>						
0.3 / -0.1	63	0.21	0.32	0.36	2.06	0.26
0.4 / -0.2	38	0.21	0.32	0.37	0.91	0.15
<i>Panel B: IVV Underpriced / SPY Overpriced</i>						
0.3 / -0.1	29	0.22	0.34	0.42	1.19	0.21
0.4 / -0.2	14	0.33	0.45	0.54	1.19	0.24

Profit summary statistics for strategies that require 0.3% (0.4%) profit for the arbitrage to be opened and are closed at a loss of 0.1% (0.2) to ensure an overall profit of 0.2%.