

# Information Asymmetry Around S&P 500 Index Changes

Youna Hong

A Thesis

in

The John Molson School of Business

Presented in Partial Fulfillment of the Requirements  
For the Degree of Master of Science in Administration at  
Concordia University  
Montreal, Quebec, Canada

December 2012

© Youna Hong, 2012

**CONCORDIA UNIVERSITY**

**School of Graduate Studies**

This is to certify that the thesis prepared

By: **Youna Hong**

Entitled: **Information Asymmetry Around S&P 500 Index Changes**

and submitted in partial fulfillment of the requirements for the degree of

**Master of Science in Administration (Finance Option)**

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Xiao Huang Chair

Dr. Sandra Betton Examiner

Dr. Nilanjan Basu Examiner

Dr. Rahul Ravi Supervisor

Approved by Dr. Harjeet S. Bhabra  
Chair of Department or Graduate Program Director

Dr. Harjeet S. Bhabra  
Dean of Faculty

Date December 5, 2012

## ABSTRACT

### Information Asymmetry Around S&P 500 Index Changes

Youna Hong

We examine the effects of S&P 500 index composition changes on the trading environment from 2001 to 2010. We find that the adverse selection cost of trading significantly decreases post addition and increases post deletion. However, the intraday price dynamics of additions to the index seem to be distinct from those of deletions from the index. The event period cumulative abnormal returns for additions are significantly associated with the change in the adverse selection cost of trading while this association is non-significant for deletions. However, the cumulative abnormal returns for deletion event periods are found to be significantly associated with the change in realized spreads. Realized spreads may be interpreted as a measure of revenue earned by the liquidity providers in the market.

## **ACKNOWLEDGEMENT**

I thank my supervisor Dr. Rahul Ravi for his insightful advice, heartfelt encouragement and support throughout this work. Through his guidance, I gained a better understanding of market microstructure, and confidence to conduct more imposing work.

I am indebted to Dr. Sandra Betton, Dr. Nilanjan Basu and Dr. Yaxuan Qi for their helpful comments. Any errors that may remain are solely my responsibility.

I dedicate this thesis to my beloved family, Marc and Ethan.

## Table of Contents

I. Introduction .....	1
2. Literature Review.....	6
2.1. On Index Changes.....	6
2.2. On Information Asymmetry.....	9
3. Data and Methods .....	10
3.1. S&P 500 Index Constituent Changes.....	10
3.2. Trades and Quotes.....	12
3.3. Decomposition of Spreads .....	13
4. Results.....	14
4.1. Descriptive Statistics.....	14
4.2. Abnormal Returns around Index Changes .....	15
4.3. Changes in Liquidity and Information Asymmetry around Index Changes .....	18
4.4. Regressions of Event Returns on Changes in Liquidity and Information Asymmetry .....	20
A. Additions.....	25
B. Deletions.....	26
4.4. Robustness Tests.....	27
5. Conclusion .....	28
References.....	30
Table 1    Sample Construction of S&P 500 Index Changes from 2001 to 2010.....	i
Table 2    Descriptive Statistics of Trading Environment around S&P 500 Changes.....	ii
Table 3    Cumulative Abnormal Returns around S&P 500 Changes .....	iii
Table 4    Comparison of Trading Variables between Pre- and Post-Index Changes .....	iv
Table 6    Effects of Liquidity, Information Asymmetry, Arbitrage Risk and Investor Recognition around Index Changes .....	vi
Table 7    Robustness Tests.....	viii

## **I. Introduction**

Extant researches have documented that companies added to the S&P 500 index exhibit price increases on average, while those deleted from the index experience price decreases. Several explanations that attempt to explain these effects may be classified into two categories based primarily on the premise of informational effect. The hypotheses that assume no informational changes – transitory price pressure (Harris and Gurel, 1986; Erwin and Miller, 1998), and long-term downward-sloping demand curves for stocks (Shleifer, 1986; Wurgler and Zhuravskaya, 2002) – and the hypotheses that address changes in information – on operating performance (Denis, McConnell, Ovtchinnikov, and Yu, 2003), earnings quality and disclosures (Platikanova, 2008), sustained liquidity effect (Beneish and Whaley, 1996; Hegde and McDermott, 2003) and investor awareness (Chen, Noronha, and Singal, 2004).

This study explores information asymmetry (as measured by the adverse selection component of the bid-ask spread) around S&P 500 revisions. Intuitively, improved disclosure and transparency post addition can either enable uninformed investors to “catch up” to informed investors (thereby reduced inter-investor information asymmetry), or benefit informed investors to gain a greater advantage through their superior processing of the increased information (thereby increased inter-investor information asymmetry). Hegde and McDermott (2003) find evidence of reduced information asymmetry post addition, while no changes in the asymmetric information component for

deletions.<sup>1</sup> Will the deletion results change with larger sample? To what extent are the changes in information asymmetry associated with the price formation around index changes? These are empirical questions with important implications. First, an answer will give us a better understanding of how information is incorporated into security prices in the market and what impact the changes in informational environment has on the trading environment. Second, it could enhance our understanding of the price effects of addition to and deletion from indices.

The sample period for this study ranges from 2001 to 2010. Consistent with the findings of Hegde and McDermott (2003), we find that quoted spreads and relative spreads decrease (increase) over the three months following addition (deletion). Thus we find evidence of a significant long-term improvement in liquidity post addition and significant long-term deterioration in liquidity post deletion.

We also find a significant decline in the adverse selection cost post addition and a corresponding increase in the realized spread. Realized spreads may be interpreted as non-informational component of the price impact. The presence of informed traders will cause market prices on average to rise after buyer-initiated trades and, similarly, fall after seller-initiated trades. Due to these adverse price changes, liquidity providers, in general, earn less than the effective spread, and the realized spread is a measure of their revenue net of losses to better-informed traders. For our sample, the changes in effective spreads following addition are found to be statistically zero in the [-60, +60] window. This implies that increased firm transparency (Platikanova, 2008) post addition, in turn, may

---

<sup>1</sup> Hegde and McDermott caution against reading too much into the results for deletions due to the small sample size.

lead to lower risk of adverse selection in the trading environment.<sup>2</sup> However, a large portion of the benefits from this change seem to be getting captured by liquidity suppliers, resulting in effective spreads remaining unchanged. We find this result somewhat surprising, since we expect that the instrument of post-addition improved liquidity to be competition among liquidity providers.<sup>3</sup>

Conversely, we find that the adverse selection cost of trading significantly increases over the three months post deletion. One possible reason for this change is perhaps the exit of index funds that, in turn, reduces non-informational trading. It can also be potentially related to a decline in the number of analysts covering the firm (Zhou, 2011) that leads to a lower amount of firm-specific information (thereby increased firm-to-investor information asymmetry).

Surprisingly, the realized spreads post deletion appear to decline significantly, resulting in narrower effective spreads, inconsistent with the increase in quoted and relative spreads. Quoted spreads are the difference between the bid and ask quotes and measure market and order flow conditions at the time of order arrival, whereas effective spreads consider trades that are executed inside the quotes, therefore, more reliably measuring the true cost of trading.<sup>4</sup> One possible explanation for the lower effective spreads is a greater level of competition among liquidity suppliers. Given the increase in

---

<sup>2</sup> Platikanova (2008) finds that discretionary accruals significantly decrease after additions, potentially improving earnings quality.

<sup>3</sup> Hendershott, Jones, and Menkveld (2011) find increased realized spreads for large-cap stocks under algorithmic trading environment, but no significant effects for small-cap stocks. To the extent that the trading in our sample period is influenced by algorithmic trading, it may be at least partially explained by this mechanism.

<sup>4</sup> Effective spreads measure how much above the midpoint price one pays on a buy order and how much below the midpoint price one receives on a sell order. Thus, effective spreads capture the magnitude of price improvement.



asymmetric information post deletion, it might be explained by the models that present an inverse relation between liquidity supply and informed order flow.<sup>5</sup>

Furthermore, we present evidence that abnormal returns following addition are positively associated with changes in adverse selection costs. This suggests that *ceteris paribus*, improvements in liquidity post addition, caused by a decline in adverse selection, should be associated with lower cumulative abnormal returns (CAR). However, to the extent that the post-addition CAR is positive, other factors than informed trading alone should affect the results. Various reasons for increased post-addition CAR have been suggested including a demand shock from indexers (Blouin, Raedy, and Shackelford, 2000), arbitrage risk (Wurgler and Zhuravskaya, 2002), a change in expected cash flows (Denis et al., 2003), an improvement in liquidity (Hegde and McDermott, 2003).

In addition, we find that post-deletion abnormal returns are significantly negatively associated with changes in realized spreads in the immediate window. These findings may imply that post-addition returns are associated with a sustained asymmetric information effect, whereas post-deletion returns are associated with a transitory liquidity effect. Presently, we do not have a compelling explanation for this pattern. These results may at least be considered as additional evidence supporting asymmetric effects between additions and deletions, consistent with Chen et al. (2004).

Finally, the results of our analysis suggest that a decrease in firm-to-investor information asymmetry post addition (Platikanova, 2008) is associated with a decrease in

---

<sup>5</sup> Huang and Stoll (1996) and DeJong et al. (1996) find a negative correlation of realized spread and adverse selection costs.

inter-investor information asymmetry. Similarly, an increase in post deletion firm-to-investor information asymmetry (Zhou, 2011) is associated with an increased inter-investor information asymmetry. This finding supports the catching-up hypothesis presented above, whereby, improved disclosure post addition enables the uninformed investors to catch up to the better-informed investors, leading to lower adverse selection risk. On the contrary, Huson and McKinnon (2003) find that reduction in firm-to-investor information asymmetry post focus-enhancing spinoff, leads to greater adverse selection risk in the market. In their study, an increase in the firm-to-investor information asymmetry creates an even more favorable environment for superiorly informed investors, thereby exacerbating adverse selection problem.

These seemingly contradictory results suggest that the relationship between firm-to-investor information asymmetry and inter-investor information asymmetry cannot be monotonic. In an attempt to explain the contradiction, Ravi and Hong (2012) propose a unimodal relation between the two types of information asymmetry. They find that the inter-investor information asymmetry increases, and then declines, as the firm-to-investor information asymmetry decreases from absolutely opaque situations to perfectly transparent situations.

We conjecture that spinoff firms are likely to be relatively opaque prior to spinoff and therefore, decreased opacity may lead to increased adverse selection. In contrast, additions to the S&P 500 are likely to be relatively large and stable firms, with greater transparency. Therefore, these firms are probably positioned on the decreasing information asymmetry side of the opacity and information asymmetry dimension, where

marginal gain of information acquisition is smaller than marginal cost of being uninformed, thereby, resulting in lower information asymmetry between investors.

The rest of this paper is organized as follows: Section 2 summarizes related literature. Section 3 describes our sample construction and research methods. Section 4 reports and discuss main results. Section 5 concludes the paper.

## **2. Literature Review**

### *2.1. On Index Changes*

Here, we provide background information of key literatures on the index changes. Shleifer (1986) and Harris and Gurel (1986) are one of the earliest works on the S&P 500 index changes, followed by Jain (1987). Their main hypotheses on the price effects of revisions to the S&P 500 index may be categorized as the downward sloping demand curve hypothesis, the price pressure hypothesis, and the information hypothesis. More recently, the investor recognition hypothesis has been proposed.<sup>6</sup>

At the onset, the price pattern around index changes garnered attention as potential evidence refuting market efficiency. The imperfect substitute hypothesis associates permanent price effects with downward sloping demand curves for stocks due to difficulty of finding perfect substitutes and, therefore, a limit-to-arbitrage case (Lynch and Mendenhall, 1997; Shleifer, 1986; Wurgler and Zhuravskaya, 2002).<sup>7</sup> In comparison, the price pressure hypothesis addresses transitory price effects accompanied by

---

<sup>6</sup> Many previous works provide a detailed survey of existing hypotheses. See, for instance, Chen et al. (2004), Elliott et al. (2006), Hegde and McDermott (2003) and Kappou et al. (2008).

<sup>7</sup> In Shleifer's study, ten days after the index change is regarded as permanent.

temporary order imbalances among various index replicators (Harris and Gurel, 1986; Elliott and Warr, 2003). Broadly speaking, the imperfect substitute hypothesis and the price pressure hypothesis can be classified together as the long-term and short-term downward sloping demand curves, differing mostly on the duration of the effect. For instance, Pruitt and Wei (1989) find that changes in institutional ownership are positively correlated to the changes in the S&P 500, augmenting both Shleifer (1986) and Harris and Gurel (1986).

In contrast with the information-free premise of these two hypotheses, the information hypothesis argues that the permanent price effects are caused by the incorporation of unexpected information on the future performance of the respective companies. Various evidences consistent with this line of research have been proposed including price increases of the supplementary index additions that are unrelated to index fund rebalancing (Jain, 1987), concurrent price changes of stocks, bonds, and options of the index inclusions and exclusions (Dhillon and Johnson, 1991), significant changes in earnings forecasts and realized earnings post-index changes (Denis, McConnell, Ovtchinnikov, and Yu, 2003), spillover effects on the companies in the same industry as the additions (Cai, 2007; Gyax and Otchere, 2010) and changes in earnings quality post additions (Platikanova, 2008). Furthermore, the permanent price impact may be driven by the permanent changes in direct trading costs, suggested by the liquidity-wealth hypothesis (Amihud and Mendelson, 1986).

Although the above hypotheses expect differences in continuation of the price effects, they all expect symmetric changes for additions and deletions. On the contrary, the investor recognition hypothesis attempts to explain asymmetric changes in prices –

permanent price increases for additions but temporary price decreases for deletions (Chen, Noronha, and Singal, 2004) using Merton's (1987) shadow cost. The shadow cost measure is pioneered by Kadlec and McConnell (1994) which finds that investor recognition plays an important role on the returns of the new listings on NYSE. As a further extension, Zhou (2011) documents differences in the price effects of subgroups across additions and deletions: first-time additions to the S&P 500 index family exhibit greater changes in investor recognition and permanent price increases, whereas migrations from other S&P indices, re-entries to the S&P 500 and deletions show transitory price effects.

Microstructural perspectives that are focused on liquidity effects include Erwin and Miller (1998) and Hegde and McDermott (2003). Erwin and Miller (1998) find a significant decrease in the spreads, associated with enhanced stock liquidity, for additions which do not have listed options. However, for optioned stocks, they failed to find a significant change in the spreads, potentially attributing the discrepancy to differences in informational efficiencies between optioned and non-optioned stocks. Hegde and McDermott (2003) present comprehensive research in market microstructural aspects, and find that additions show improved liquidity mainly due to lower transacting costs, and deletions exhibit deteriorated liquidity.

Predatory trading that intends to exploit index funds from announcement to effective dates has been documented by Beneish and Whaley (1996; 2002). Research that examines trading of index funds include Blume and Edelen (2004), Petajisto (2010) and Green and Jame (2011). Green and Jame (2011) finds evidence of strategic trading by index funds and differentiated price effects depending on the source of liquidity

provision. However, their findings those stocks held widely by small- and mid-cap funds experience smaller return can also be attributed to investor recognition or information effects due to a joint hypothesis problem.

So far, empirical tests of various hypotheses have been mixed; for example, Kappou et al. (2008) do not find support for investor recognition and price pressure, while Elliott et al. (2006) claims that investor recognition is a primary factor. Recent works using firm characteristics, industry level data or markets other than equities (Kasch and Sarkar, 2011; Qi and Sui, 2007; Gyax and Otchere, 2010) increasingly add support for the information hypothesis and our finding based on transaction level data is also consistent with informational effects.

## *2.2. On Information Asymmetry*

Following Huson and MacKinnon (2003) who analyzed the effect of corporate spinoffs on information asymmetry, we propose two converse views about whether asymmetric information among investors will increase or decrease following the index changes.

First, information asymmetry between investors can decrease for additions and increase for deletions. More public information will reduce the value of private information, which Hasbrouck (1991) suggests as a key advantage of informed traders. Therefore, following index additions (deletions), more (less) public information and investor awareness may reduce (increase) inter-investor information asymmetry.

In contrast, index additions (deletions) can increase (decrease) information asymmetry due to an increased activity of informed traders, as the added (deleted) stocks

now have a wider (limited) following, thereby, potentially enlarging the benefits of being informed. In addition, superior processing of increased disclosure by informed traders can also exacerbate information asymmetry. This line of research where better public information complements informed traders' private information follows the models of Kim and Verrecchia (1994) and Lundholm (1991).

### **3. Data and Methods**

#### *3.1. S&P 500 Index Constituent Changes*

Unlike some indices that mechanically follow predetermined rebalancing rules, the S&P 500 index is maintained by the S&P Index Committee, who monitor and evaluate the representativeness of the index against the U.S. equities market. Generally, index changes are triggered by deletions that are related to mergers and acquisitions of the index constituents. Other corporate events such as restructuring, liquidation or bankruptcies also initiate deletions. The remaining deletions are caused by a substantial violation of one or more of the criteria for index inclusion (S&P Indices, 2011).

The guidelines for index additions are public information, although the identification of candidate firms is not disclosed. According to the S&P index methodology, index additions are selected based on the following criteria: U.S. company, market capitalization, public float, financial viability, adequate liquidity and reasonable

price, sector representation, and company type. Inevitably, there is a sufficient amount of judgement involved in the changes of the S&P 500 components.<sup>8</sup>

[Insert Table 1]

We obtain a list of the index changes from 2001 to 2010 from Standard and Poor's, and verify each reason of the changes by manually examining the official public release statements. Our sample period is a post-decimalization era and includes market structure changes in NYSE after 2005. The initial sample consists of 256 additions and 256 deletions, as the index changes are concurrent to maintain 500 constituents. To be included in our analysis, a firm should not be engaged in contemporaneous corporate events. After filtering deletions that are related to mergers and acquisitions, which are the most common reasons of index changes, the initial sample consists of 212 additions and 65 deletions. Following Chen et al. (2004), we include only samples with return data of at least 60 trading days prior to the announcement date and 90 trading days after the effective date on CRSP and trades and quotes data on TAQ for the same window. We also eliminate other confounding events such as bankruptcy, leveraged buyouts, spin-offs, delisting, relocation to a foreign country, and cosmetic changes such as name changes by searching news using Bloomberg Professional. Further, we screen out stocks with multiple additions and deletions history.

Our final sample is 178 additions and 51 deletions, yielding a fairly large sample size of deletions than that of most other studies. Previous studies are generally confined to additions only, due to data collection difficulties in deletions. Table 1 shows that the

---

<sup>8</sup> Most studies assume that additions and deletions cannot be predicted in advance. On the contrary, Geppert et al. (2010) claims that deletions from the S&P 500 can be predicted. This supports our thesis that announcement of deletions are not associated with changes in informational content of trading.



annual frequency of the S&P 500 changes is not uniformly distributed, as the years 2008 and 2009 have higher representations of the changes. Since the years of 2007 and 2008 correspond to abnormal market environments, we examine the effect of these years using dummy variables later in our regression analysis.

Since 1989, Standard and Poor's have pre-notified index changes with on average 5 trading days between the announcement and effective dates. Both announcement and index changes take place after the close of trading. We conduct our analysis on a period surrounding announcement, as any information effects from announcement should be incorporated immediately in an efficient capital market.

### *3.2. Trades and Quotes*

We extract trades and quotes data of the S&P 500 changes, 90 days prior to the announcement dates and 90 days post the effective dates. Following Hegde and McDermott (2003) we define our event periods from sixty days prior to announcements to sixty days post announcement [-60, +60]. We extract trade and quote data from the NYSE TAQ database for the event period.

Trades data are filtered through the standard criteria<sup>9</sup> and contemporaneous trades are combined using a volume-weighted average price (VWAP). Similarly, we filter quotes data and create NBBO quote per time stamp.<sup>10</sup> We match trades and quotes using

---

<sup>9</sup> All trades with a correction indicator other than 0 or 1, conditions A, C, D, N, O, R, or Z, non-positive size or price, and trades recorded before opening time (9:30) or after the closing time (16:00) of the market are excluded.

<sup>10</sup> All quotes with non-positive bid size or price, ask size or price, negative bid-ask spreads where offer price is less than bid price, quotes for which the quoted spread is greater than 20% of the quote midpoint when the quote midpoint is greater than \$10, or for which the quoted spread is greater than \$2 when the quote midpoint is less than \$10, and quotes with condition 5, 7, 8, 9, 11, 13, 14, 15, 16, 17, 19, 20, 27, 28, 29 are excluded.

Lee and Ready (1991)'s algorithm – applying a tick test and a quote test – with a one-second lag, and assign trade direction indicators (buyer-initiated trades = 1, seller-initiated = -1, other = 0) to all quotes.

### 3.3. Decomposition of Spreads

Liquidity effects of the S&P 500 index changes are documented by Hegde and McDermott (2003), who use quoted spread, depth, effective spreads, and adverse selection cost of trading measures based on the models of Glosten and Harris (1988), Foster and Viswanathan (1993), and Madhavan, Richardson, and Roomans (1997). In our analysis, we use the Hendershott et al. (2011) adaptation of the Glosten and Harris (1988) model to estimate the permanent price impact (adverse selection parameter) and the transient price impact measures.

We measure liquidity and information asymmetry using effective half-spreads, 5-minute realized spreads, and 5-minute price impacts normalized on the prevailing quote midpoint, respectively. The proportional effective half-spread ( $ESPRD_{it}$ ) is the signed difference between the midpoints of the bid and ask quotes and the trade price.

$$ESPRD_{it} = \frac{q_{it}(p_{it} - m_{it})}{m_{it}},$$

where  $q_{it}$  is a trade direction indicator from Lee and Ready (1991) for stock  $i$  at the time  $t$ ,  $p_{it}$  is the trade price, and  $m_{it}$  is the prevailing quote midpoint. For each stock, we measure quoted and effective spreads for each reported transaction and calculate a share-weighted average across all trades that day.

We decompose spreads into two parts: the 5-minute realized spread ( $RSPRD_{it}$ ) is a proxy for liquidity cost or revenue to liquidity providers under the assumption that the dealers can cover their positions at the quote midpoint 5 minutes after the trade. If a quote is not available at the exact 5-minute interval, the closest quote post 5 minutes is substituted.

$$RSPRD_{it} = \frac{q_{it}(p_{it} - m_{i,t+5min})}{m_{it}},$$

where  $m_{i,t+5min}$  is the prevailing quote midpoint 5 minutes after the trade at  $t$ . The adverse selection or losses due to informed liquidity demanders is defined as

$$ADVSEL_{it} = \frac{q_{it}(m_{i,t+5min} - m_{it})}{m_{it}}.$$

Therefore, there exists an identity for the effective spread as the sum of the realized spread and the adverse selection.

$$ESPRD_{it} = RSPRD_{it} + ADVSEL_{it}.$$

## 4. Results

### 4.1. Descriptive Statistics

Table 2 contains summary statistics on trading variables – number of trades, realized spread, adverse selection, effective spread, trade size (= number of shares traded), and turnover (= trading volume normalized on shares outstanding). Spread variables are winsorized at 99% and 1% to reduce influences by top and bottom outliers. Throughout the sample period, trading has substantially grown; the number of trades,

trade size, and turnover are far larger in the sub-period of 2006 to 2010 than in the period of 2001-2005.

In general, additions are more liquid than deletions: mean (median) realized spread and adverse selection of additions are 0.0868 (0.0220) and 0.0384 (0.0278) bps, while mean (median) realized spread and adverse selection of deletions are 0.1888 (0.0380) and 0.1513 (0.0656) bps. Further, additions seem to be larger than deletions, given the lower mean (median) trade size of additions of 2.8 (1.8) mn shares versus the mean (median) trade size of deletions of 5.1 (3.0) mn shares, despite their similar number of trades between additions and deletions.

Between the sub-periods, there seems to be no visible differences in the median adverse selection cost for both additions and deletions. However, realized spread, in general, has widened in the period of 2006-2010, implying a potential market structural change in the recent trading environment.

[Insert Table 2]

#### *4.2. Abnormal Returns around Index Changes*

Previous studies have thoroughly documented abnormal returns surrounding the S&P 500 changes. However, there exists little consensus across them when it comes to explaining either the magnitude or the duration of the price effects around the events. For instance, Shleifer (1986) and Harris and Gurel (1986) examine similar sample periods and reach opposing conclusions on whether the effects are transient or permanent. The variations in the results are somewhat inevitable due to variations in selecting event study

parameters. In this study, we only address the direction of the price reactions, without concern for whether the effect is temporary or permanent.

We perform a daily event study based on the Fama-French and the momentum factor model using the CRSP equally-weighted index (Fama and French, 1993; Carhart, 1997).<sup>11</sup> The results from the market model are qualitatively similar. To the extent that the additional factors (size, market-to-book ratio, and momentum) control for latent price effects, our choice of benchmark is more stringent. Our event study parameters are as follows: 255 trading days for the estimation period, 15 days between the end of estimation period and the start of the event window, and 3 days of minimum non-missing returns for estimation. As the composition changes are announced after the market closes, we set the announcement day (AD) as the day after the announcement.

[Insert Table 3]

The basic abnormal return results, omitted here for brevity, are consistent with previous findings; abnormal returns of 2.98% for additions and -4.00% for deletions on the announcement day, significant at 1%, that are partially reversed in subsequent days. Compared to the data until 2000 from Chen et al. (2004), the magnitudes of the returns are smaller in our samples, potentially indicating that the trading pattern has changed. From the Table 3, there is no substantial evidence of abnormal price changes prior to the announcement date for index changes, as the cumulative abnormal returns (CAR) in [-10,-2] and [-5,-2] are not significant at conventional levels. The CAR results suggest that the effect of potential information leakages during the specified periods is not significant.

---

<sup>11</sup> In comparison, Chen et al. (2004) use the S&P 500 index as a benchmark.

Table 3 shows that the price gains post additions are more than offset by subsequent declines, and vice versa for deletions. After roughly a month and a half post announcement (30 trading days post announcement), the performance of stocks included into the S&P 500 seem to be worse than the performance of stocks excluded from the index. The CAR for additions is -3.44% (generalized sign z-test = -1.82), while the CAR for deletion is 1.39% ( $z = -0.93$ ) thirty trading days post announcement. These results are not driven by outliers, as 104 out of 178 added stocks present negative returns and 29 out of 51 deleted stocks exhibit positive returns during the window of  $[-2,+30]$ . Although the 30-day cumulative abnormal return for deletions is statistically insignificant, the degree of price effects in deletion is more pronounced than that of additions in most of the examined event windows.

While at first glance these results might seem to contradict the assertions of the extant literature, we find that it is somewhat consistent with the findings of some of the more recent studies such as Zhou (2011). The study finds that pure deletions have a larger price increase (11.96%) than the gain by pure additions (7.13%) using samples up to 2008.<sup>12</sup> The perverse price reactions between additions and deletions are also documented by fund managers. For instance, a 2005 report from Hussman Funds says that “Given such dour sentiment for stocks removed from the S&P 500, the group has shown surprisingly strong returns, consistently outperforming the shares of companies that have been added to the index.”<sup>13</sup>

---

<sup>12</sup> For definitions of subgroups such as pure additions and pure deletions, see Zhou (2011): “Companies added to the S&P index family, for the first time” are pure additions, and “companies dropped from the S&P index family altogether” are pure deletions.

<sup>13</sup> See “Misfit Stocks: Deletions from the S&P 500 regularly outperform new additions”, William Hester, March 2005 (<http://hussman.net/rsi/misfitstocks.htm>)

A possible explanation for this result may be found in the literature dealing with strategic trading. While pseudo-arbitrageurs attempt to front-run index funds (Beneish and Whaley, 1996), the index funds themselves spread out their rebalancing over a large window in an attempt to avoid large price impact (Green and Jame, 2011). The proliferation of algorithmic trading in the market and the availability of high-frequency trading platforms would also potentially change the post-2001 results. Other potential reasons may include behavioural responses that create undervaluation and, in turn, profit opportunities for neglected stocks.

#### *4.3. Changes in Liquidity and Information Asymmetry around Index Changes*

Chen et. al (2004) argue that the price effect around addition is persistent, while the effect around deletion is transient. Building on their results, we conjecture that the observed abnormal returns are potentially related to the activities of informed traders in the market. An increase in disclosure of information around addition can possibly lead to a higher level of informed trading, or, on the contrary, it can reduce the information asymmetry due to a diminished informational advantage by informed traders, as more private information becomes public.

Similarly, the deletion event leads to a reduced number of analysts following the firm (Zhou, 2011), potentially changing the informational environment. In one end, the adverse selection cost might rise due to increased firm-to-investor information asymmetry, on the other end, the adverse selection cost might decline, as the diminished release of information can potentially make relatively a large fraction of investors equally uninformed and, therefore, no informational distinction among them. Hegde and

McDermott (2003) present some evidence that the adverse selection post addition declines. However, the deletion result is still an empirical issue.

Table 4 reports that quoted spreads and relative spreads decrease (increase) over the three months following addition (deletion). Hence, there is evidence of a significant long-term improvement in liquidity post addition and significant long-term deterioration in liquidity post deletion.

We find a significant drop in adverse selection post addition and an increase in the realized spread. This is as suggested by theoretical market microstructural models, which predicts that liquidity supply and informed order flow are inversely related. The realized spread, which is a measure of the non-permanent price impact of trade, captures the liquidity provider's revenue after loss to informed traders. In the presence of informed order flow, liquidity supplier increases the spread in order to recoup the losses that may occur when trading against better-informed traders, leading to a negative relation between realized spread and adverse selection.

For our sample, the changes in effective spreads following addition are statistically zero. This implies that increased firm transparency (Platikanova, 2008) post addition, in turn, may lead to lower risk of adverse selection in the trading environment. However, a large portion of the benefits from this change seem to be exploited by liquidity suppliers, resulting in effective spreads remaining unchanged. This result is somewhat surprising, since we expect that the instrument of post-addition improved liquidity to be competition among liquidity providers.



Conversely, we find that the adverse selection cost of trading significantly increases over the three months post deletion. One possible reason for this change is perhaps the exit of index funds that, in turn, reduces non-informational trading. It can also be potentially related to a decline in the number of analysts covering the firm (Zhou, 2011) that leads to a lower amount of firm-specific information (thereby increased firm-to-investor information asymmetry).

Surprisingly, the realized spreads post deletion appear to decline significantly, resulting in narrower effective spreads, inconsistent with the increase in quoted and relative spreads. Quoted spreads are the difference between the bid and ask quotes and measure market and order flow conditions at the time of order arrival, whereas effective spreads consider trades that are executed inside the quotes, therefore, more reliably measuring the true cost of trading. One possible explanation for the lower effective spreads is a greater level of competition among liquidity suppliers. Given the increase in asymmetric information post deletion, it might also be explained by the models that present an inverse relation between liquidity supply and informed order flow.

#### *4.4. Regressions of Event Returns on Changes in Liquidity and Information Asymmetry*

The results so far have been univariate. In this section, we further analyse the association between the changes in adverse selection risks and realized spreads around index changes and the corresponding CAR around the events. We control for the effects from some of the previously documented associations, such as arbitrage risk (ArbRisk) and shadow costs (dShadow). If the index changes were a mainly liquidity driven event, we expect to see a significant relation between the non-permanent price impact parameter (realized spread) and the CAR, but a non-significant association between changes in the

permanent price impact measure (adverse selection). Otherwise, the index changes may convey informational effect, and, therefore, a permanent effect. An adverse selection cost of trading is a measure of illiquidity. Therefore, changes in information asymmetry should be positively related to returns. Signs of realized spread variable remain an empirical issue.

We decompose the effective spread (ESPRD) into the realized spread (RSPRD) and the adverse selection (ADVSEL), which represents transitory price impact or revenue that compensates liquidity providers and permanent price impact or informational loss of liquidity provider due to informed trading, respectively. Following Hendershott et al. (2011), RSPRD proxies for liquidity effect and ADVSEL proxies for information asymmetry. Our basic analytical framework is:

$$CAR = \alpha + \beta_1 dRSPRD + \beta_2 dADVSEL + \epsilon,$$

$$\text{where } dRSPRD = PostRSPRD - PreRSPRD$$

$$dADVSEL = PostADVSEL - PreADVSEL$$

In specifying the regressions, we follow Chen et al. (2004), which examine the relation between abnormal returns and changes in Merton's shadow cost. As in McConnell, Servaes, and Lins (2008), which measure changes in insider ownership on changes in firm value, the method of using changes is preferred to using the level of each variable by enabling to circumvent potential endogeneity between the level of liquidity, information asymmetry and returns.

We measure changes in liquidity and information asymmetry by the mean differences between windows of [-60, -10] and [+10, +60] in RSPRD, ADVSEL, and shadow cost variables. These methods are adopted from Hegde and McDermott (2003), and Chen et al. (2004). By dropping the immediate event intervals, we intend to avoid potential contamination effects from the index changes. Using the mean differences without omitting the intervals yields qualitatively similar results with a stronger significance.

Further, we control for the arbitrage risk effect (Wurgler and Zhuravskaya, 2002) and the investor recognition effect (Chen et al., 2004). In previous studies, arbitrage risk and investor recognition are found to have generally significant influences on abnormal returns (Elliott et al., 2006; Green and Jame, 2011), while other control factors such as size, age, NYSE/Nasdaq dummies, and changes in analysts' forecasts are not consistently significant in multivariate testing (Elliott et al., 2006). Our expanded model including control variable are:

$$CAR = \alpha + \beta_1 dRSPRD + \beta_2 dADVSEL + \beta_3 ArbRisk + \beta_4 dShadow + \epsilon.$$

We proxy for arbitrage risk by residual standard deviation and investor recognition by Merton's shadow cost, respectively. We construct a variable that measures arbitrage risk in the spirit of Wurgler and Zhuravskaya (2002), which propose that high-arbitrage-risk stocks have steeper demand curves than low-arbitrage-risk stocks, and, therefore, a positive sign of the ArbRisk coefficient. Arbitrage risk is measured as the variance of the residuals of the Fama-French and the momentum factor model,

$$R_{it} - R_{ft} = \alpha + \beta_{1i}(R_{mt} - R_{ft}) + \beta_{2i}SMB + \beta_{3i}HML + \beta_{4i}UMD,$$

where  $R_{mt}$  is the return on the CRSP equally-weighted index,  $R_{ft}$  is the one-month Treasury Bill rates,  $SMB$  is the small-minus-big return,  $HML$  is the high-minus-low return, and  $UMD$  is the Carhart momentum factor.<sup>14</sup> The parameters are estimated over a [-365, -20] trading day window. In Wurgler and Zhuravskaya (2002), two arbitrage-risk measures, A1 and A2, are suggested: A1 is the variance of the residuals of the excess return market model regression estimated over a [-365, -20] day window. A2 is the residual variance of the excess returns on three industry-, size-, and book-to-market-matched close substitute stocks. We only examine A1 in our analysis, since A1 and A2 have a correlation of 0.98 (Wurgler and Zhuravskaya, 2002), and our abnormal returns are based on the four-factor model that separately incorporates size and book-to-market ratio effects.

We proxy for investor awareness (recognition) using Merton's shadow cost. Merton (1987) suggested that investors only trade in the stocks of which they are aware, thereby holding incompletely diversified portfolios. The difference between the higher returns for less than fully diversified investors than that for completely diversified investors is Merton's shadow cost.

Following Kadlec and McConnell (1994) and Chen et al. (2004), we define shadow cost as:

$$Shadow\ Cost = \frac{Residual\ Std\ Dev}{S\&P\ 500\ Market\ Cap} \times \frac{Firm\ Size}{No.\ of\ Shareholders}$$

---

<sup>14</sup> Wurgler and Zhuravskaya (2002) use a market model residual with the CRSP value-weighted index. We modify the benchmark to be consistent with our abnormal return measures. We verify that the correlations among the market model residual, the Fama-French three-factor model residual, and the four-factor model are above 0.99, justifying the use of the multi-factor model residual.

where residual standard deviation is measured as the standard deviation of the difference between the firm's return and the S&P 500 return in the 252-day period before the announcement date and in the 252-day period after the effective date. Firm size, measured by the market value of the firm's equity, and the S&P 500 market capitalization are calculated on the announcement date. The number of shareholders for the pre-event period is measured as close as possible prior to the announcement date, and the number of shareholders for the post-event period is as close as possible after the effective date. Price and number of shares outstanding are from CRSP and the number of shareholders is from Compustat.

Prior to regressions, we examine correlations among variables. Trade variables – trading volume, size, number of trades – are correlated at a higher level to each other, but our key variables of interests such as RSPRD, ADVSEL, ArbRisk, and shadow costs are not seriously correlated. Among the key variables, ADVSEL and ArbRisk are most highly correlated at 0.1539. This result intuitively makes sense, as stocks with higher idiosyncratic risk are likely to be more obscure and should have higher presence of information asymmetry that rewards informed trading.

The regression results are reported in Table 6. All our results are based on heteroscedasticity-consistent robust standard errors. Multicollinearities are not a big concern to affect the stability and size of the standard errors, as the variance inflation factors (VIFs, unreported) are less than 2 for additions and less than 4 for deletions. The elevated VIFs for the spreads in deletions suggest a potential collinearity is a bigger issue in deletions, while they remain below the conventionally regarded as a critical level of

10. The univariate results with spreads-only tests are analogous to the multivariate results; hence, we primarily discuss the multivariate results.

[Insert Table 6]

#### *A. Additions*

The intercepts (Table 6) of additions shows significance throughout -10 days to +10 days of the window, implying that the cumulative abnormal returns are statistically significant, similar to the event study results reported in Section 4.1. We find that changes in adverse selection (dADVSEL) are influential during the [-2,+1], [-2,+5], and [-2,+10] windows in explaining the abnormal returns (t-statistics = 1.69, 2.25, and 3.02). The relation suggests that the increasing inter-investor information asymmetry during these periods is positively correlated to the cumulative abnormal returns. The result of the [-2,+1] window captures an immediate price reaction, while those of [-2,+5] and [-2,+10] correspond to the price responses up to the effective dates and including a period when index funds are expected to strategically adjust their positions (Green and Jame, 2011). The result presents an evidence of informed trading that is significantly associated with the price reactions immediately around the S&P 500 index additions.

The results of ArbRisk and dShadow are inconsistent with the limited arbitrage risk and investor recognition hypotheses. According to the arbitrage risk hypotheses, stocks with high idiosyncratic risk should experience larger abnormal returns on inclusion into the S&P 500 index. Although ArbRisk exhibit a significance at 10% for CAR[-5,-2] and CAR[-2,+30], the negative signs for additions contradicts the expectations that stocks with larger arbitrage risk should associate with larger price

changes. We may consider the following reconciliations: First, Wurgler and Zhuravskaya (2002) use additions between 1976 and 1989, before the preannouncement of index changes. Although this setting may be regarded as a clean test that can remove confounding effects from potential front running of investors other than index funds, the sample period is undeniably outdated. The trading environment has changed significantly since then. Similarly,  $dShadow$  is inconsistent with the investor recognition hypothesis: Although the lower shadow cost suggests improved investor awareness that results in higher abnormal returns, the signs change after  $CAR[-2,+5]$  and their statistical significance are very small. We omit the  $CAR[-2,+30]$  and  $CAR[-2,+60]$  results for brevity, but they are qualitatively similar to those of shorter intervals at a weaker significance.

### *B. Deletions*

The regression results of deletions are distinct from those of additions. There are evidences that improved liquidity (= negative  $dRSPRD$ ) is an important factor to the abnormal returns (t-statistics of  $CAR[-2,+1] = -2.77$ ) at 1% level, however, the effect is short-lived. Unlike additions, we do not find evidence that changes in adverse selection is associated with the event returns. Arbitrage risk and shadow cost are also insignificant in general and their signs are not as the same expected by extant hypotheses.

Our results suggest that information asymmetry has a significant association with  $CAR$  for additions, but not for deletions. Although the realized spreads are negatively related, the coefficients are not significant for any of the event windows. This discrepancy to the Hegde and McDermott (2003) results may be at least partly attributed

to the differences in the sample periods. We conjecture that informational effects from announcement of deletions are limited, since investors who are familiar with the S&P 500 index may already have incorporated any fundamental information into the prices. These results suggest that trading around the announcement of deletions is influenced by latent variables other than the spreads, arbitrage risk, and shadow cost. However, an elaboration of the determinants of deletion returns should be left for future study.

#### 4.4. Robustness Tests

In this section, we report the robustness test results. First, we conduct analysis using the adverse selection measure by Glosten and Harris (1988, hereafter GH). The GH model defines the adverse selection, and the direct cost of transacting such as the inventory holding, and the order-processing components as a linear function of transaction volume:

$$\Delta P_t = c_0 \Delta I_t + c_1 \Delta V_t + z_0 I_t + z_1 I_t V_t + \epsilon_t$$

where  $P_t$  is a transaction price,  $I_t$  is a trade direction indicator of 1 (buyer-initiated trades), and -1 (seller-initiated trades), following the Lee and Ready (1991) procedure,  $V_t$  is a trading volume at a time  $t$ . The information asymmetry component is measured as percentage of the bid-ask spread:

$$GH = \frac{2(z_0 + z_1 \bar{V})}{2(z_0 + z_1 \bar{V}) + 2(c_0 + c_1 \bar{V})} \times 100$$

where  $\bar{V}$  is the average transaction volume,  $2(z_0 + z_1 \bar{V})$  is the adverse selection component, and  $2(c_0 + c_1 \bar{V})$  is the transitory component of the bid-ask spread.



[Insert Table 7]

Using GH as a measure of information asymmetry, we find that information asymmetry slightly increases in the index changes during the window of  $[-2,+1]$  (t-statistics = 1.70 for additions and 1.76 for deletions, respectively). However, in the windows of  $[-2,+5]$  and  $[-2,+10]$ , the GH measures do not show a clear pattern. Given that the GH measure is calculated on an annual basis, the power of the measure may be weaker.

Additionally, we test the changes in trading volume normalized by outstanding shares (dVol), and the effect of the financial crisis (yDum as 1 if the year is 2007 or 2008, and 0 otherwise). Both covariates do not appear significant in explaining the cumulative abnormal returns around index changes, and our conclusions from Section 4.3 remain intact. Hence, the general pattern of adverse selection and liquidity remain the same after controlling for trading volume and the year-effect.

## **5. Conclusion**

This study examines asymmetric information between investors around S&P 500 revisions using a recent sample from 2001 to 2010. We find that a significant sustained improvement in liquidity, as measured by quoted spreads and relative spreads, post addition and the converse relation post deletion. We find evidence of a decrease in inter-investor asymmetry post addition, potentially led by an improved dissemination of information from a firm to the market. This result is consistent with the theoretical predictions by Diamond (1985) and Hakansson (1977). Furthermore, the decline in the

adverse selection cost over the subsequent three months is significantly positively associated with the event period returns of additions, exhibiting the price formation process around additions. In contrast, we do not find evidence that changes in inter-investor asymmetric information are related to event window returns of deletions. This asymmetric result may be broadly consistent with the findings by Chen et al. (2004) who find asymmetric price behaviors and investor recognition between additions and deletions.

This study helps better understand the dynamics of two types of information asymmetry – one from a firm to investor, and the other between investors, presents evidence of the role of adverse selection in index changes, and extends the previous results from market microstructural aspects (Hegde and McDermott, 2003) with an alternative measure of liquidity and adverse selection cost.

For future research, we first recognize that much has been changed in the era of algorithmic trading. Our evidence indicates that price effects around S&P 500 index changes in the recent period, which is coincident to the prevalence of algorithmic trading, are dissimilar to the findings from prior literature. Therefore, a study of index changes that are associated with algorithmic trading will be useful to gain an understanding of modern trading environment. Another potential research topic from our study is an examination of how liquidity and information asymmetry changes in the predatory trading context. Although index changes do not yield distressed trading per se, fore-known trading by index funds can be regarded as an approximation for predatory trading situation.

## References

- Amihud, Y., & Mendelson, H. (1986). Asset Pricing and the Bid-Ask Spread. *Journal of Financial Economics*, 17, 223-249.
- Beneish, M. D., & Whaley, R. E. (1996). An anatomy of the "S&P Game": The effects of changing the rules. *Journal of Finance*, 51(5), 1909-1930.
- Beneish, M. D., & Whaley, R. E. (2002). S&P 500 index replacements: A new game in town. *Journal of Portfolio Management*, 29(1), 51-60.
- Blume, M. E., & Edelen, R. M. (2004). S&P 500 indexers, tracking error, and liquidity. *Journal of Portfolio Management*, 37-46.
- Cai, J. (2007). What's in the news? Information content of S&P 500 additions. *Financial Management*, 36(3), 113-124.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.
- Chen, H., Noronha, G., & Singal, V. (2004). The price response to S&P 500 index additions and deletions: Evidence of asymmetry and a new explanation. *Journal of Finance*, 50(4), 1901-1930.
- Denis, D. K., McConnell, J. J., Ovtchinnikov, A. V., & Yu, Y. (2003). S&P 500 index additions and earnings expectations. *Journal of Finance*, 58(5), 1821-1840.
- Dhillon, U., & Johnson, H. (1991). Changes in the Standard and Poor's 500 list. *Journal of Business*, 64(1), 75-85.
- Elliott, W. B., & Warr, R. S. (2003). Price pressure on the NYSE and Nasdaq: Evidence from S&P 500 index changes. *Financial Management*, 85-99.
- Elliott, W. B., Van Ness, B. F., Walker, M. D., & Warr, R. S. (2006). What drives the S&P 500 inclusion effect? An analytical survey. *Financial Management*, 35(4), 31-48.
- Erwin, G. R., & Miller, J. M. (1998). The liquidity effects associated with addition of a stock to the S&P 500 index: Evidence from bid/ask spreads. *Financial Review*, 33(1), 131-146.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3-56.
- Foster, F. D., & Viswanathan, S. (1993). Variation in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *Journal of Finance*, 48(1), 187-211.

- Geppert, J. M., Ivanov, S. I., & Karels, G. V. (2010). Analysis of the probability of deletion of S&P 500 companies: Survival analysis and neural networks approach. *The Quarterly Review of Economics and Finance*, 50(2), 191-201.
- Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21, 123-142.
- Green, T. C., & Jame, R. (2011). Strategic trading by index funds and liquidity provision around S&P 500 index additions. *Journal of Financial Markets*, 14, 605-624.
- Gygax, A. F., & Otchere, I. (2010). Index composition changes and the cost of incumbency. *Journal of Banking & Finance*, 34, 2500-2509.
- Harris, L., & Gurel, E. (1986). Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. *Journal of Finance*, 41(4), 815-829.
- Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance*, 46, 179-207.
- Hegde, S. P., & McDermott, J. B. (2003). The liquidity effects of revisions to the S&P 500 index: an empirical analysis. *Journal of Financial Markets*, 6, 413-459.
- Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *Journal of Finance*, 66(1), 1-33.
- Huson, M. R., & MacKinnon, G. (2003). Corporate spinoffs and information asymmetry between investors. *Journal of Corporate Finance*, 9, 481-503.
- Jain, P. C. (1987). The effect on stock price of inclusion in or exclusion from the S&P 500. *Financial Analysts Journal*, 43(1), 58-65.
- Kadlec, G. B., & McConnell, J. J. (1994). The effect of market segmentation and illiquidity on asset prices: Evidence from exchange listings. *Journal of Finance*, 49(2), 611-636.
- Kappou, K., Brooks, C., & Ward, C. W. (2008). A re-examination of the index effect: Gambling on additions to and deletions from the S&P 500's 'gold seal'. *Research in International Business and Finance*, 22, 325-350.
- Kasch, M., & Sarkar, A. (2011). Comovement revisited. *Federal Reserve Bank of New York Staff Reports No. 484*.
- Kim, O., & Verrecchia, R. E. (2004). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, 17, 41-67.
- Lee, C. M., & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance*, 46(2), 733-746.

- Lundholm, R. J. (1991). Public signals and the equilibrium allocation of private information. *Journal of Accounting Research*, 29(2), 322-349.
- Lynch, A. W., & Mendenhall, R. R. (1997). New evidence on stock price effects associated with changes in the S&P 500 index. *Journal of Business*, 70(3), 351-383.
- Madhavan, A. (2003). The Russell reconstitution effect. *Financial Analysts Journal*, 59(4), 51-64.
- Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. *Review of Financial Studies*, 10(4), 1035-1064.
- McConnell, J. J., Servaes, H., & Lins, K. V. (2008). Changes in insider ownership and changes in the market. *Journal of Corporate Finance*, 14, 92-106.
- Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42(3), 483-510.
- Petajisto, A. (2010). The index premium and its hidden cost for index funds. *Journal of Empirical Finance*, 18(2), 271-288.
- Pruitt, S. W., & Wei, K. J. (1989). Institutional ownership and changes in the S&P 500. *Journal of Finance*, 44(2), 509-513.
- Qi, R., & Sui, L. (2007). The addition and deletion effects of the Standard & Poor's 500 index on option markets. Working paper.
- Ravi, R., & Hong, Y. (2012). Firm Opacity and Financial Market Information Asymmetry. *Working Paper*.
- S&P Indices. (2011, Sep.). S&P U.S. indices methodology.
- Shankar, S. G., & Miller, J. M. (2006). Market reaction to changes in the S&P SmallCap 600 index. *Financial Review*, 41(3), 339-360.
- Shleifer, A. (1986). Do demand curves for stocks slope down? *Journal of Finance*, 41(3), 579-590.
- Wurgler, J., & Zhuravskaya, E. (2002). Does arbitrage flatten demand curves for stocks? *Journal of Business*, 75(4), 583-608.
- Zhou, H. (2011). Asymmetric changes in stock prices and investor recognition around revisions to the S&P 500 index. *Financial Analysts Journal*, 67(1), 72-84.

Table 1 Sample Construction of S&P 500 Index Changes from 2001 to 2010

This table describes a list of changes to the S&P 500 index from 2001 to 2010, after removing confounding events such as mergers and acquisitions, bankruptcies, and other corporate restructuring, as well as multiple inclusions of and exclusions from the index. Two additions whose effective dates occurred in January 2001 are included to additions in 2001.

Year	Initial samples		Final samples		% proportion of changes	
	additions	deletions	additions	deletions		
2001	31	31	20	5	11.2	9.8
2002	24	24	13	5	7.3	9.8
2003	10	10	7	0	3.9	0.0
2004	20	20	11	6	6.2	11.8
2005	20	20	12	2	6.7	3.9
2006	31	31	24	6	13.5	11.8
2007	40	40	29	3	16.3	5.9
2008	34	34	30	10	16.9	19.6
2009	29	29	18	12	10.1	23.5
2010	17	17	14	2	7.9	3.9
Total	256	256	178	51	100	100

Table 2 Descriptive Statistics of Trading Environment around S&P 500 Changes

This table summarizes the trading environment of the S&P 500 index changes from 2001 to 2010. The statistics are estimated over -60 to +60 event window (with 0 being the announcement day). Number of trades is the number of trades; price is the actual trade price per share; RSPRD is the 5-minute realized spread; ADVSEL is the information asymmetry component of the bid-ask spread; ESPRD is the effective spread following Hendershott et al. (2011); trade size is the number of shares traded; turnover is the trade volume normalized by outstanding shares.

2001-2010	A. Additions (N = 21625)			B. Deletions (N = 7276)		
	Mean	Std Dev	Median	Mean	Std Dev	Median
Number of trades	4,339.95	3,557.86	3,472.00	4,574.25	4,052.20	3,567.50
Price	59.74	73.20	43.80	18.68	16.06	16.63
RSPRD ( $\times 10^2$ )	0.0868	0.2018	0.0220	0.1888	0.5455	0.0380
ADVSEL ( $\times 10^2$ )	0.0384	0.0522	0.0278	0.1513	0.6640	0.0656
ESPRD ( $\times 10^2$ )	0.1275	0.2277	0.0533	0.3464	1.0209	0.1180
Trade size	2,884,957	3,683,217	1,750,569	5,063,367	7,875,360	2,953,950
Turnover	16.32	22.55	10.27	22.54	27.46	15.49
2001 - 2005	A. Additions (N = 7978)			B. Deletions (N = 2285)		
Number of trades	2,489.47	2,843.30	1,527.00	1,220.79	1,220.74	913.00
Price	42.49	17.50	39.22	19.73	11.19	19.25
RSPRD ( $\times 10^2$ )	0.0396	0.0966	0.0196	0.1155	0.5215	0.0238
ADVSEL ( $\times 10^2$ )	0.0399	0.0451	0.0294	0.2481	1.1418	0.0658
ESPRD ( $\times 10^2$ )	0.0816	0.1188	0.0552	0.3676	1.5863	0.0967
Trade size	2,382,773	3,958,412	1,204,650	2,539,352	8,177,827	432,500
Turnover	11.88	19.38	6.58	11.50	19.67	6.14
2006 - 2010	A. Additions (N = 13647)			B. Deletions (N = 4991)		
Number of trades	5,421.73	3,487.23	4,588.00	6,109.55	3,968.64	5,137.00
Price	69.83	89.64	46.09	18.20	17.83	12.73
RSPRD ( $\times 10^2$ )	0.1144	0.2388	0.0241	0.2223	0.5529	0.0514
ADVSEL ( $\times 10^2$ )	0.0376	0.0559	0.0268	0.1070	0.1997	0.0656
ESPRD ( $\times 10^2$ )	0.1543	0.2683	0.0515	0.3367	0.6062	0.1420
Trade size	3,178,533	3,479,101	2,157,632	6,218,921	7,453,756	4,411,136
Turnover	18.92	23.84	12.47	27.59	28.99	20.93

Table 3 Cumulative Abnormal Returns around S&P 500 Changes

Abnormal returns are measured compared to the Fama-French-Momentum factor model using the CRSP equally-weighted index as a benchmark. The event study parameters are as follows: 255 trading days for the estimation period, 15 days between the end of estimation period and the start of the event window, and 3 days of minimum non-missing returns for estimation. As the composition changes are announced after the market closes, the day after the announcement is set as the announcement day (AD). Mean abnormal returns are in percent. Positive: Negative describes the proportion of returns that are positive and negative, respectively. The symbols <sup>\*\*\*</sup>, <sup>\*\*</sup>, and <sup>\*</sup> denote statistical significance at the 1%, 5%, and 10%, respectively, using a one-tail test. The significance of the proportions similarly corresponds to the significance using the generalized sign test.

Days relative to the AD	A. Additions (N = 178)			B. Deletions (N = 51)		
	Cumulative abnormal return	Positive : Negative	Generalized sign Z	Cumulative abnormal return	Positive : Negative	Generalized sign Z
[-10,-2]	0.09	82:96	-0.623	-7.00	20:31 <sup>*</sup>	-1.490 <sup>*</sup>
[-5,-2]	-0.13	89:89	0.427	-3.48	20:31 <sup>*</sup>	-1.490 <sup>*</sup>
[-2,+1]	2.81	148:30 <sup>***</sup>	9.276 <sup>***</sup>	-6.91	10:41 <sup>***</sup>	-4.291 <sup>***</sup>
[-2,+5]	2.02	118:60 <sup>***</sup>	4.777 <sup>***</sup>	-6.44	12:39 <sup>***</sup>	-3.731 <sup>***</sup>
[-2,+10]	0.71	99:79 <sup>**</sup>	1.927 <sup>**</sup>	-1.98	18:33 <sup>**</sup>	-2.050 <sup>**</sup>
[-2,+20]	-1.40	82:96	-0.623	1.35	25:26	-0.090
[-2,+30]	-3.44	74:104 <sup>**</sup>	-1.823 <sup>**</sup>	1.39	22:29	-0.930



Table 4 Comparison of Trading Variables between Pre- and Post-Index Changes

This table presents changes in the spreads, trading volume, and prices around index changes. The changes in the variables are measured over -60 to -1 and +1 to +60 trading days, and over -60 to -10 and +10 to +60 trading days. Day zero represents the announcement day. RSPRD, ADVSEL, and ESPRD are relative spread, adverse selection, and effective spread, respectively. QSPRD is quoted spread, or the difference between the bid and ask quotes. RelSPRD is the relative spread, or QSPRD divided by the price. nVol is the trading volume, nTrades is the number of trades, price is the transaction price, and size is the trading size. The symbols <sup>\*\*\*</sup>, <sup>\*\*</sup>, and <sup>\*</sup> denote statistical significance at the 1%, 5%, and 10%, respectively.

Panel A: [-60, -1] vs. [+1, +60]

	A. Additions		B. Deletions	
	Post - Pre	t-values	Post - Pre	t-values
RSPRD	0.00003 <sup>**</sup>	2.15	-0.0005 <sup>***</sup>	-11.23
ADVSEL	-0.00002 <sup>***</sup>	-5.59	0.0002 <sup>***</sup>	6.98
ESPRD	0.00002	1.13	-0.0003 <sup>***</sup>	-5.46
QSPRD	-0.0106 <sup>***</sup>	-3.41	0.0613 <sup>***</sup>	23.81
RelSPRD	-0.0003 <sup>***</sup>	-7.35	0.0031 <sup>***</sup>	20.5
nVol	4.8124 <sup>***</sup>	37.63	4.5967 <sup>***</sup>	20.36
nTrades	841.2 <sup>***</sup>	66.85	-331.1 <sup>***</sup>	-14.69
Price	0.7889 <sup>***</sup>	11.81	-2.8137 <sup>***</sup>	-39.52
Size	832702 <sup>***</sup>	44.63	873513 <sup>***</sup>	17.22

Panel B: [-60, -10] vs. [+10, +60]

	A. Additions		B. Deletions	
	Post - Pre	t-values	Post - Pre	t-values
RSPRD	0.0001 <sup>**</sup>	4.75	-0.0005 <sup>***</sup>	-10.73
ADVSEL	-0.00001 <sup>***</sup>	-3.19	0.0002 <sup>***</sup>	4.54
ESPRD	0.00007 <sup>***</sup>	4.17	-0.0004 <sup>***</sup>	-5.86
QSPRD	-0.0166 <sup>***</sup>	-4.74	0.0679 <sup>***</sup>	22.63
RelSPRD	-0.0004 <sup>***</sup>	-8.08	0.0033 <sup>***</sup>	20.67
nVol	2.1023 <sup>***</sup>	27.26	0.7606 <sup>***</sup>	4.97
nTrades	728.7 <sup>***</sup>	55.88	-622.1 <sup>***</sup>	-25.26
Price	0.6356 <sup>***</sup>	8.14	-3.1686 <sup>***</sup>	-37.62
Size	514831 <sup>***</sup>	34.6	114708 <sup>***</sup>	2.72

Table 5 Correlations among Variables

This table describes the Pearson correlation coefficients among the independent variables under the null hypothesis of no correlation, measured cross-sectionally and across years from 2001 to 2010, around the S&P 500 index changes. nTrades is the number of trades; price is the actual trade price per share; RSPRD is the 5-minute realized spread; ADVSEL is the information asymmetry component of the bid-ask spread; ESPRD is the effective spread from Hendershott and (2011); size is the number of shares traded; nVol is the trade volume normalized by outstanding shares; Arbitrage risk is measured as the variance of the residuals of the Fama-French and the momentum factor model. Shadow cost is defined as (Residual Standard Deviation)/(S&P 500 Market Capitalization) × (Firm Size)/(Number of Shareholders).

	nTrades	PRICE	RSPRD	ADVSEL	ESPRD	SIZE	nVol	ArbRisk	preShadow
nTrades									
Price	0.0714								
RSPRD	0.2238	-0.0686							
ADVSEL	0.0067	-0.0738	0.3004						
ESPRD	0.1521	-0.0883	0.8230	0.7774					
Size	0.6732	-0.0822	0.2697	0.2114	0.3026				
nVol	0.5151	0.0377	0.2150	0.1074	0.2078	0.6458			
ArbRisk	0.0499	-0.0935	0.0269	0.1539	0.1066	0.2321	0.1746		
preShadow	0.0988	0.1506	0.0018	-0.0163	-0.0087	0.0239	0.1140	0.1776	
postShadow	0.1343	0.1979	0.0136	-0.0206	-0.0040	0.0260	0.1415	0.1407	0.9123

Table 6 Effects of Liquidity, Information Asymmetry, Arbitrage Risk and Investor Recognition around Index Changes

The table contains univariate and multivariate tests of the S&P 500 index changes, using the model  $CAR = \alpha + \beta_1 dRSPRD + \beta_2 dADVSEL + \beta_3 ArbRisk + \beta_4 dShadow + \epsilon$ , where  $dRSPRD = PostRSPRD - PreRSPRD$ , and  $dADVSEL$  and  $dShadow$  are analogously defined. Changes in liquidity, information asymmetry, and investor recognition are measured by the mean differences between windows of [-60, -10] and [+10, +60] in RSPRD, ADVSEL, shadow cost variables. RSPRD is the 5-minute realized spread, and ADVSEL is the information asymmetry component of the bid-ask spread from Hendershott and (2011); Arbitrage risk is measured as the variance of the residuals of the Fama-French and the momentum factor model. Shadow cost is defined as (Residual Standard Deviation)/(S&P 500 Market Capitalization)  $\times$  (Firm Size)/(Number of Shareholders). All results are based on heteroscedasticity-consistent robust standard errors. Chen et al. (2004) report  $dShadow$ , deflating at  $10^{-9}$ . The values in parentheses are t-statistics. The symbols <sup>\*\*\*</sup>, <sup>\*\*</sup>, and <sup>\*</sup> denote statistical significance at the 1%, 5%, and 10%, respectively.

Panel A: Additions												
	CAR[-10,-2]		CAR[-5,-2]		CAR[-2,+1]		CAR[-2,+5]		CAR[-2,+10]		CAR[-2,+20]	
Intercept	0.002	0.021 <sup>*</sup>	-0.001	0.017 <sup>**</sup>	0.029 <sup>***</sup>	0.033 <sup>***</sup>	0.021 <sup>***</sup>	0.025 <sup>***</sup>	0.009	0.013 <sup>*</sup>	-0.013	0.004
	(0.41)	(1.85)	(-0.23)	(2.01)	(7.83)	(6.38)	(4.79)	(4.54)	(1.63)	(1.74)	(-1.66)	(0.32)
dRSPRD	-7.629	-9.249	-3.271	-4.869	-1.469	-1.813	-1.397	-1.756	-5.402	-5.639	-9.938	-11.510
	(-1.30)	(-1.47)	(-0.79)	(-1.17)	(-0.38)	(-0.45)	(-0.29)	(-0.36)	(-1.03)	(-1.06)	(-1.15)	(-1.27)
dADVSEL	61.625	69.878 <sup>*</sup>	6.102	14.839	42.009 <sup>*</sup>	43.1440 <sup>*</sup>	54.990 <sup>**</sup>	57.173 <sup>**</sup>	89.624 <sup>***</sup>	90.053 <sup>***</sup>	33.210	37.503
	(1.57)	(1.72)	(0.27)	(0.66)	(1.72)	(1.69)	(2.21)	(2.25)	(3.05)	(3.02)	(0.83)	(0.89)
ArbRisk		-32.960		-31.7981 <sup>*</sup>		-7.588		-7.454		-7.408		-30.139
		(-1.33)		(-1.73)		(-0.99)		(-0.83)		(-0.66)		(-1.14)
dShadow		-56.128		-61.276		1.959		-30.350		4.947		60.150
		(-0.43)		(-0.65)		(0.05)		(-0.44)		(0.07)		(0.57)
R-square	0.020	0.0798	0.004	0.134	0.030	0.038	0.036	0.040	0.054	0.056	0.008	0.041
N	175	175	175	175	175	175	175	175	175	175	175	175

Panel B: Deletions

	CAR[-10,-2]		CAR[-5,-2]		CAR[-2,+1]		CAR[-2,+5]		CAR[-2,+10]		CAR[-2,+20]	
Intercept	-0.051***	-0.075***	-0.026*	-0.041*	-0.059***	-0.061*	-0.055**	-0.065**	-0.019	-0.018	0.018	0.009
	(-2.95)	(-3.21)	(-1.69)	(-1.79)	(-2.74)	(-1.90)	(-2.64)	(-2.43)	(-0.61)	(-0.43)	(0.46)	(0.16)
dRSPRD	-31.772***	-28.683***	-11.580	-10.456	-26.395***	-26.059***	-7.024	-4.489	11.587	13.524	5.076	6.067
	(-3.22)	(-3.07)	(-1.01)	(-0.92)	(-2.68)	(-2.77)	(-0.35)	(-0.23)	(0.63)	(0.76)	(0.24)	(0.29)
dADVSEL	3.720	0.384	0.340	-1.785	6.097	5.723	-2.538	-3.932	-5.996	-5.878	-4.952	-6.247
	(0.73)	(0.08)	(0.05)	(-0.27)	(1.16)	(1.18)	(-0.29)	(-0.50)	(-0.76)	(-0.83)	(-0.53)	(-0.73)
ArbRisk		31.061**		23.876		3.539		6.943		-11.017		13.060
		(2.07)		(1.63)		(0.18)		(0.49)		(-0.33)		(0.29)
dShadow		-7045.12***		-229.79		-737.54		-9214.23		-10066.00		-1683.86
		(-3.36)		(-0.12)		(-0.34)		(-1.53)		(-1.35)		(-0.23)
R-square	0.337	0.389	0.091	0.110	0.125	0.126	0.072	0.116	0.010	0.035	0.007	0.009
N	51	51	51	51	51	51	51	51	51	51	51	51

Table 7 Robustness Tests

This table contains robustness tests of the S&P 500 index changes, using the model from Table 6 as a base. GH for information asymmetry is measured following Glosten and Harris (1988). dVol is the change from pre-event normalized trading volume to post-event normalized trading volume. All results are based on heteroscedasticity-consistent robust standard errors. The values in parentheses are t-statistics. The symbols \*\*\*, \*\*, and \* denote statistical significance at the 1%, 5%, and 10%, respectively.

Panel A: Additions									
	CAR[-2,+1]			CAR[-2,+5]			CAR[-2,+10]		
Intercept	0.033*** (6.39)	0.012 (0.87)	0.031*** (5.39)	0.027*** (4.91)	0.009 (0.55)	0.023*** (3.59)	0.017** (2.36)	0.004 (0.23)	0.015* (1.76)
dRSPRD	-1.895 (-0.43)		-2.310 (-0.52)	-0.518 (-0.10)		-1.200 (-0.24)	-3.256 (-0.55)		-3.516 (-0.59)
dADVSEL	42.959* (1.68)		42.564* (1.68)	59.987** (2.31)		59.339** (2.32)	95.468*** (3.13)		95.221*** (3.10)
GH		0.0006* (1.70)			0.0005 (0.91)			0.0002 (0.35)	
ArbRisk	-7.601 (-0.99)	-2.924 (-0.37)	-6.610 (-0.85)	-7.248 (-0.82)	-3.518 (-0.36)	-5.621 (-0.62)	-7.012 (-0.66)	-4.060 (-0.35)	-6.393 (-0.59)
dShadow	0.282 (0.01)	37.919 (0.82)	-1.052 (-0.02)	-4.861 (-0.06)	17.124 (0.22)	-7.050 (-0.09)	53.998 (0.67)	69.334 (0.94)	53.167 (0.66)
dVol	0.0001 (0.1)		0.0000 (-0.02)	-0.0001 (-0.85)		-0.0010 (-0.96)	-0.0017 (-1.23)		-0.0017 (-1.26)
yDum			0.005 (0.58)			0.009 (0.93)			0.003 (0.28)
R-square	0.038	0.030	0.040	0.051	0.012	0.055	0.083	0.005	0.083
N	175	175	175	175	175	175	175	175	175

Panel B: Deletions

	CAR[-2,+1]			CAR[-2,+5]			CAR[-2,+10]		
Intercept	-0.062*	-0.129	-0.033***	-0.0643**	-0.094	-0.056**	-0.018	0.015	-0.068**
	(-1.90)	(-1.86)	(-3.07)	(-2.38)	(-1.42)	(-2.63)	(-0.42)	(0.16)	(-2.56)
dRSPRD	-26.049***		-14.805	-4.496		-1.190	13.518		-5.717
	(-2.74)		(-1.25)	(-0.23)		(-0.07)	(0.76)		(-0.31)
dADVSEL	4.589		0.366	-3.151		-4.392	-5.134		2.090
	(0.83)		(0.05)	(-0.55)		(-0.86)	(-0.80)		(0.25)
GH		0.003*			0.001			-0.001	
		(1.76)			(0.92)			(-0.48)	
ArbRisk	4.508	-4.322	-4.212	6.275	-13.781	3.711	-11.653	-20.111	3.263
	(0.23)	(-0.14)	(-0.37)	(0.41)	(-0.53)	(0.28)	(-0.35)	(-0.62)	(0.13)
dShadow	-1050.965	-2851.117	1201.294	-8998.274	-10380.000	-8335.958	-9860.120	-9487.149	-13713*
	(-0.47)	(-1.19)	(0.35)	(-1.53)	(-1.57)	(-1.35)	(-1.31)	(-1.43)	(-1.69)
yDum	0.0009		0.0007	-0.0006		-0.0007	-0.0006		-0.0002
	(0.50)		(0.38)	(-0.23)		(-0.26)	(-0.17)		(-0.06)
dfc			-0.089			-0.026			0.153
			(-0.90)			(-0.47)			(1.37)
R-square	0.127	0.042	0.160	0.116	0.073	0.119	0.035	0.027	0.089
N	51	51	51	51	51	51	51	51	51