SCHOOL OF ECONOMICS AND MANAGEMENT

LUND UNIVERSITY

Does mispricing explain returns following addition to the S&P 500?

by

James Nunn – 841130 6999

Giacomo Reghelin – 961129 T379

May, 2020

Master's Programme in Finance

Supervisor: Dr. Jens Forssbaeck

Abstract

This study investigates the relationships between mispricing and returns following addition to the S&P 500. We find evidence that the index premium is demand-driven, resulting in mispricing that is subsequently exploited by rational investors. We find that index premia are temporary, undergoing a full reversal within 20 days. Using a novel approach to estimating mispricing, we find that post-addition returns are negatively related to mispricing. Furthermore, the average degree of mispricing decreases following addition, lending support to the theory that increased scrutiny due to addition to the S&P 500 aids investors in identifying and exploiting pricing inefficiencies.

Keywords: *Index premium; S&P 500; Mispricing; Market efficiency*

Table of Contents

List of Tables

1 Introduction

Since its origins in the Sharpe (1964) and Lintner (1965) capital asset pricing model (CAPM), fund indexing and passive investment in market indices have grown in popularity, fuelled by findings that actively managed equity funds fail to beat market indices. S&P Dow Jones Indices (2020b) report that over the 15 years to December 2019, 90.5% of all U.S. large-cap funds underperform the S&P 500 Index.

This study investigates the effect of addition to the S&P 500. The S&P 500 measures the market performance of the stocks of the 500 largest companies listed on eligible exchanges in the U.S., weighted by float-adjusted market capitalisation. It is widely considered to be a proxy for the U.S. equity market (S&P Dow Jones Indices, 2020a) and is commonly treated as such by passive investors. Constituent changes are made by the S&P Index Committee based on transparent eligibility criteria. Consistent with the Index Committee's stated objective of avoiding constituent turnover (S&P Dow Jones Indices, 2020a), eligibility criteria apply to addition to the index as opposed to ongoing membership. Companies are not necessarily deleted for breaching eligibility criteria, although a company that substantially violates one or more of the criteria may be deleted at the Index Committee's discretion. Deletions commonly occur when a company is involved in a merger, acquisition, or significant restructuring such that it no longer meets the eligibility criteria. Constituent changes take effect from the *effective addition day*.

Research has consistently found evidence for the existence of an *index premium*: positive (negative) abnormal returns associated with addition to (deletion from) key indices (Shleifer, 1986; Harris & Gurel, 1986; Chen, Noronha & Singal, 2004; Petajisto, 2011). *Demand-based* theories attribute the index premium to the demand shock due to the actions of index-tracking investors when an index undergoes a constituent change. Index-tracking investors hold a portfolio of an index's constituent stocks weighted according to the stock's weight in the index. Since the S&P 500 has a fixed constituent count of 500, deleted stocks are replaced with a suitable candidate. When the index undergoes a constituent change, index-tracking investors sell the deleted stock and purchase the added stock, resulting in a negative (positive) demand shock for the deleted (added) stock. Demand-based theories assume that constituent changes do not convey information to the market and that the index premium is due to downward-sloping demand curves. This contradicts the assumption of horizontal demand curves under the

Efficient Markets Hypothesis (EMH), under which, in the absence of information effects, any price change represents a market inefficiency that will be exploited and eliminated via the actions of *rational investors*. Conversely, *information-based* theories of the index premium propose that constituent changes convey information regarding a firm's future prospects to the market.

This study investigates an extension of the *short-term* demand-based theory of the index premium. If short-run demand curves are downward-sloping and *addition premia* are indeed driven by demand from index-tracking investors, according to the EMH, the resulting market inefficiencies will be exploited and eliminated via the actions of rational investors. Thus, index premia are expected to be temporary. Furthermore, rational investors will sell (buy) overvalued (undervalued) stocks driving prices to fundamental values. Thus, returns following addition are expected to be negatively related to *mispricing*.

We use additions to the S&P 500 from 1995 to January 2020 and adopt a novel approach to estimating mispricing, to investigate whether returns subsequent to addition are negatively related to mispricing. The timing of any price correction depends on the market's ability to identify market inefficiencies. As such, we investigate the relationships between mispricing and both addition premia and post-addition returns. Indeed, the actions of rational investors may reasonably be expected not to occur until after the effective addition day when demand from index-tracking investors peaks, in which case any price correction would be captured by post-addition returns. We further hypothesise that increased market scrutiny and information availability upon addition to the S&P 500 will aid investors in identifying market inefficiencies, causing the degree of mispricing to decrease.

This study comprises three primary applications of empirical techniques. The *primary regressions* investigate the relationships between mispricing and an added stock's returns following addition to the index, including addition premia and post-addition returns. An *event study* estimates addition premia and post-addition returns, providing the dependent variables for the primary regressions. The *auxiliary regressions* provide a measure of a stock's mispricing: the explanatory variables of interest in the primary regressions.

We confirm the existence of a temporary index premium that undergoes a complete reversal in the 20 days following addition. This is consistent with downward-sloping short-run demand curves and inconsistent with downward-sloping long-run demand curves and information-based theories, which predict that index premia should be persistent. We find that post-addition returns are negatively related to the degree of mispricing following addition, supporting the hypothesis that the index premium is demand-driven, representing a departure from market efficiency that is subsequently eliminated by rational investors. We find that index premia are independent of mispricing prior to addition, which is consistent with our hypothesis that any price correction will not occur until after addition and the corresponding demand shock. It is also consistent with the downward-sloping short-run demand curve hypothesis, under which index premia are determined by the magnitude of the demand shock due to addition to the index and the elasticities of supply and demand. We find that the degree of mispricing decreases for added firms following addition, consistent with the hypothesis that addition to the index results in increased market scrutiny and information availability, aiding investors in identifying market inefficiencies and reducing the degree of mispricing.

The remainder of this paper is divided into four sections. Section 2 reviews the existing literature, develops the research question and outlines how this study contributes to the existing literature. Section 3 details methodology including data collection and empirical techniques. Section 4 presents and interprets the results of the empirical methodologies employed. Section 5 concludes with a summary of key findings and a discussion of limitations and potential areas of further research.

2 Literature Review and Research Question

The S&P 500 measures the market performance of the stocks of the 500 largest companies listed on eligible exchanges in the U.S. and is widely considered to be a proxy for the U.S. equity market (S&P Dow Jones Indices, 2020a). It has therefore been the subject of a large volume of research into the effects of constituent changes for indices, although researchers have also investigated, among others, the Dow Jones Industrial Average (DJIA) (Beneish & Gardner, 1995), the Russell 2000 (Petajisto, 2011; Cai & Houge, 2008) and the FTSE 100 (Fernandes & Mergulhão, 2016; Mase, 2007).

Researchers have found evidence for both permanent (Shleifer, 1986; Dhillon & Johnson, 1991; Beneish & Whaley, 1996; Lynch & Mendenhall, 1997; Wurgler & Zhuravskaya, 2002; Chen, Noronha & Singal, 2004) and temporary (Harris & Gurel, 1986; Vespro, 2006; Patel & Welch; 2017) price increases upon addition to the S&P 500. Research has also found evidence that the initial stock price reaction comprises permanent and temporary components with index premia undergoing a partial reversal (Beneish & Whaley, 1996; Lynch & Mendenhall, 1997; Petajisto, 2011; Ravi & Hong, 2015). Researchers have also found evidence for negative abnormal returns associated with deletion (Lynch & Mendenhall, 1997; Chen, Noronha & Singal, 2004; Petajisto, 2011; Ravi & Hong, 2015). These results are not exclusive to the S&P 500, with Beneish and Gardner (1995) finding evidence for a deletion effect for the DJIA and Petajisto (2011) finding both addition and deletion effects for the Russell 2000 Index.

Theories attempting to explain the index premium fall under two broad categories: *demand-* and *information-based*.

2.1 Demand-Based Theories

Demand-based theories attribute the index premium to the demand shock due to the actions of index-tracking investors in response to constituent changes (Shleifer, 1986; Harris & Gurel, 1986). According to financial theory, the fundamental value of a company's stock may be calculated by discounting expected free cash flows at an appropriate risk-adjusted discount rate. According to the EMH, stock prices reflect all available information and should only change due to the revelation of new information pertaining to a firm's expected free cash flows and discount rate. In the absence of information effects, any price change represents a market inefficiency, which will be eliminated via the actions of rational investors, driving prices to fundamental values resulting in flat demand curves.

Researchers have tested the assumption of flat demand curves by investigating stock price reactions to trades of large blocks of shares, finding positive (negative) reactions to purchases (sales) (Mikkelson & Partch, 1985; Scholes, 1972), consistent with downward-sloping demand curves. However, these findings are also consistent with the information hypothesis, which proposes that the sale or purchase of a large block of shares signals information to the market, resulting in the revision of a stock's fundamental value.

Shleifer (1986) and Harris and Gurel (1986) interpret the positive abnormal returns associated with addition to the S&P 500 as evidence that demand curves for stocks are downward-sloping. They attribute the positive abnormal returns to the positive demand shock created by the mechanical buying of index-tracking investors when a stock is added to the S&P 500. Petajisto (2011) reports that about 10% of a company's shares outstanding are purchased by index-tracking investors upon addition to the S&P 500. Similarly, deletion from an index results in selling by index-tracking investors and a negative demand shock. If demand curves are indeed flat as assumed under the efficient market hypothesis, in the absence of information effects, demand shocks caused by changes in index composition should not impact stock prices.

It has been proposed that announcement of addition to (deletion from) the S&P 500 conveys positive (negative) information to the market, however Shleifer (1986) argues that the informational value of inclusion in an index is unlikely to be material on the basis that the purpose of an index is to measure the market performance of constituent stocks, not to provide a prediction of future performance. He further reasons that addition of a firm with a lower S&P credit rating conveys more positive information about the firm's future prospects compared to a firm with a higher rating, which should result in a higher stock price reaction, however they do not find that stock price reactions to addition are negatively correlated with a firm's debt quality. Harris and Gurel (1986) argue that changes in index composition should not reveal new information about expected returns since they are based on publicly available information and transparent criteria.

While Harris and Gurel (1986) find that the index premium is temporary, Shleifer (1986) finds that it is permanent, resulting in alternative explanations for the index premium.

2.1.1 Downward-Sloping Long-Run Demand Curve Hypothesis

Shleifer (1986) argues that long-run demand curves are downward sloping because for an index-tracking investor, there is no close substitute for the stock of an added firm, resulting in inelastic demand. Subsequent studies have found evidence supporting Shleifer's (1986) findings of persistent addition premia (Jain, 1987; Dhillon & Johnson, 1991; Hegde & McDermott, 2003; Chen, Noronha & Singal, 2004).

2.1.2 Downward-Sloping Short-Run Demand Curve Hypothesis

Harris and Gurel (1986) argue that since there are no substitutes for stocks that are added to the S&P 500 from the perspective of an index-tracking investor, demand curves may not be perfectly elastic in the short-term but that Scholes' (1972) theory that stocks have perfect substitutes and elastic demand may hold in the longer-term. Indeed, they find that the price increase upon addition to the S&P 500 is temporary. They argue that the index premium needs to exist to compensate investors who accommodate demand shocks by trading opposite index-tracking investors, for the transaction costs and portfolio risk that they bear. Since addition to the index increases demand from index fund managers only around the effective addition date, the price impact is expected to be temporary. These findings are corroborated by Vespro (2006).

2.1.3 Upward-Sloping Supply Curve

The demand-based theories inherently assume that supply curves are upward-sloping. Schnitzler (2018) confirms this, noting that 58% of stocks added to the S&P 500 are simultaneously deleted from the S&P MidCap 400 Index. The S&P 400 comprises stocks of the 400 next largest companies in the U.S. behind the constituents of the S&P 500. When a stock from the S&P 400 is added to the S&P 500, it is simultaneously deleted from the S&P 400. Investors who track the S&P 400 sell their positions at the same time that investors tracking the S&P 500 buy. Stocks that are constituents of the S&P 400 prior to addition to the S&P 500 therefore have more elastic supply curves resulting in a smaller addition premium, which is indeed what Schnitzler (2018) finds.

2.2 Information-Based Theories

Dhillon and Johnson (1991) and Jain (1987) propose that index constituent changes convey information pertaining to a firm’s future prospects. Denis, McConnell, Ovtchinnikov and Yu (2003) find that addition results in un upwards revision in analysts’ earnings forecasts, suggesting that the S&P Index Committee possesses superior information or analytical skills. Platinakova (2008) attributes the index premium to higher quality accounting and reporting of financial information upon addition to the S&P 500, resulting in a price increase due to decreased information risk.

2.2.1 Liquidity Hypothesis

Amihud and Mendelson (1986) find that assets with higher bid-ask spreads (lower liquidity and higher trading costs) yield higher expected returns to compensate investors for the higher trading cost. The liquidity hypothesis of index premia argues that addition to an index, particularly one as ubiquitous as the S&P 500, raises a company’s profile, increases trading volume and liquidity, decreases the bid-ask spread and trading cost, reduces the company’s cost of equity and increases value. The price impact of addition is therefore expected to persist at least as long as the stock remains a constituent of the index. Indeed, Dhillon and Johnson (1991) find that addition to the S&P 500 results in a permanent increase in trading volume (liquidity). While Beneish and Gardner (1995) fail to find a stock price impact of addition to the DJIA, they find other evidence for the liquidity hypothesis, including a negative price reaction and a decline in both trading volume and quantity of available information associated with deletion from the index. Furthermore, they find that the size of the stock price reaction for deletions is positively related to the increase in transaction costs (proxied by the bid-ask spread) associated with deletion. They also find that increased trading volume associated with addition to the index is persistent, in line with the liquidity hypothesis, but that the decreased bid-ask spread is not. However, Hegde and McDermott (2003), using a more precise measure of liquidity, find that improvements in liquidity costs persist for at least three months following addition to the S&P 500, a finding that is confirmed by Ravi and Hong (2015).

2.2.2 Investor Awareness Hypothesis

Building on Merton's (1987) theory of market segmentation, Chen, Noronha and Singal (2004) advance an investor awareness hypothesis based on information asymmetry, which suggests that the enhanced corporate profile enjoyed by constituents of the S&P 500 makes it cheaper and easier to raise capital. Improved access to capital and increased investment result in increased expected cash flows, a lower cost of capital and a higher valuation. They find a permanent price increase for additions and a non-persistent price decrease for deletions. They argue that while addition to the S&P 500 raises a company's profile, a firm's profile does not immediately suffer due to deletion. Denis et al. (2003) further argue that a price increase may be justified if increased investor awareness leads to stricter monitoring of management and reduced agency costs. Ravi and Hong (2015) find that information asymmetry decreases when a company is added to the index and that variation in addition premia is explained by the quantity of information available to the market.

2.3 Mispricing

The EMH proposes that stock prices reflect all available information and that any inefficiencies will be identified and exploited by rational investors. Empirical evidence for the EMH is mixed. Kendall (1953), Roberts (1959) and Cootner (1964) find evidence supporting the EMH, however Summers (1986) argues that evidence supporting the EMH does not necessarily imply that asset prices reflect fundamental values. He argues that inefficiencies can arise and persist in the absence of investor awareness, with subsequent studies finding that stock prices are not always efficiently priced (De Bondt & Thaler, 1985, 1987; Keynes, 1936; Shiller, 1979; Shiller, 1981). Identifying and exploiting market inefficiencies represents a major area of focus in academic research and industry alike. Previous research has attempted to use mispricing to explain the actions of investors and subsequent returns. According to the EMH, rational investors sell (buy) overvalued (undervalued) stocks, resulting in negative (positive) returns until they are efficiently priced. Researchers have investigated whether mispricing explains returns (Frankel & Lee, 1998; Penman & Sougiannis, 1998; Lee, Myers & Swaminathan, 1999; Chang, Luo & Ren, 2013), however this assumes that the market is able to correctly identify market inefficiencies.

Subsequent research has focused on corporate actions including stock repurchases, equity issues and mergers and acquisitions (M&A) to investigate the relationship between mispricing and returns. Announcement of such corporate actions conveys information to the market regarding a stock's fundamental value, alerting the market to inefficiencies. D'Mello and Shroff (2000) find that stocks that undertake stock repurchases are, on average, undervalued. Announcement of a repurchase signals to the market that the stock is undervalued, resulting in in a positive stock price reaction. Dong, Hirshleifer and Teoh (2012) find that overvalued companies are more likely to issue equity, consistent with findings that equity issues are associated with a negative stock price reaction (Myers & Majluf, 1984; Asquith & Mullins, 1986). Rhodes-Kropf, Robinson and Viswanathan (2005), Rhodes-Kropf and Viswanathan, (2004) and Shleifer and Vishny (2003) find that companies that are undervalued are more likely to be acquired and to experience a positive stock price reaction (Andrade, Mitchell & Stafford, 2001), while overvalued acquirers are more likely to offer stock as consideration, resulting in a negative stock price reaction (Andrade, Mitchell & Stafford, 2001).

A number of techniques are used in academia and industry to estimate a firm's fundamental value. These can be divided into two broad categories: *present value approaches* and *relative valuation*. Discounted cash flow (DCF) methods are commonly used in industry, while Ohlson's (1991, 1995) residual income model (RIM) is prevalent in literature. Frankel and Lee (1998) investigate whether mispricing estimated using the RIM explains returns, while Warr, Elliott, Koëter-Kant and Öztekin (2012) investigate the implications of mispricing for how rapidly highly leveraged companies reduce leverage to target levels. They find that overvalued firms decrease leverage to target levels faster than less overvalued firms, consistent with overvalued companies' proclivity to issue equity.

Relative valuation is popular in industry and the academic literature. Price-to-earnings (P/E) is particularly prevalent in industry (Pinto, Robinson & Stowe, 2019). Research commonly relies on relative valuation to capture the market's beliefs (Doukas, Kim & Pantzalis, 2010; Plenborg & Pimentel, 2016). Researchers further justify the use of relative valuation on the basis that multiples are a viable substitute for fundamental valuation (Liu, Nissim & Thomas, 2002; Plenborg & Pimentel, 2016), with Liu, Nissim and Thomas (2002) finding that pricing errors calculated using relative valuation are smaller than those calculated using present value models.

Bernström (2014) argues that relative valuation techniques employed in industry can be improved upon by conditioning valuation multiples on key value drivers. Even though the market commonly uses relative valuation by applying an industry average multiple to the company of interest, this technique is not used in isolation and certainly not by the market as a whole. It is therefore not representative of either fundamental values or market beliefs. He further reasons that because multiples are tools used to identify value, value is driven by free cash flow and the choice of multiple does not affect underlying cash flows, that any multiple used in conjunction with appropriate conditioning variables and peers should result in the same valuation as a correctly-applied present value model. Rhodes-Kropf, Robinson and Viswanathan (2005) employ a similar approach to investigate the relationship between mispricing and the actions of acquirers in M&A, conditioning market value of equity on financial fundamentals to estimate mispricing using market-to-book (M/B). Chang, Luo and Ren (2013) employ this approach to investigate the relationship between mispricing and returns.

2.4 Research Question

Demand-based theories attribute the index premium to the demand shock due to the actions of index-tracking investors when an index undergoes a constituent change (Shleifer, 1986; Harris & Gurel, 1986). Downward-sloping demand curves and upward-sloping supply curves mean that when a stock is added to an index, increased demand from index-tracking investors results in a price increase. Under demand-based theories, the magnitudes of index premia are determined by the magnitude of the demand shock and the elasticities of supply and demand.

Previous research into demand-based theories of the index premium have been limited to investigating the persistence of returns. Shleifer (1986) finds that addition premia are persistent, supporting the downward-sloping long-run demand curve hypothesis, however Harris and Gurel (1986), Vespro (2006) and Patel and Welch (2017) find the index premium to be temporary, suggesting that demand curves are downward-sloping in the short-term. This study contributes to the literature by extending the analysis to investigate whether mispricing explains returns subsequent to addition. Furthermore, we adopt an approach to estimating mispricing, espoused by Bernström (2014), which, while being novel considering the existing literature, potentially better captures a stock's fundamental value and the market's beliefs.

We find that addition premia are temporary, undergoing a complete reversal in the 20 days following addition, in support of the downward-sloping short-run demand curve hypothesis. In the absence of information signalling as argued by Shleifer (1986) and Harris and Gurel (1986), any stock price reaction to addition to the index represents a departure from market efficiency which, according to the EMH, will be exploited and eliminated by rational investors who sell (buy) overvalued (undervalued) stocks driving prices to fundamental values. Thus, returns subsequent to addition are expected to be negatively related to mispricing.

The timing of any price correction depends on the market's ability to identify market inefficiencies. As such, we investigate the relationships between mispricing and both addition premia and post-addition returns. Investigation of post-addition returns represents a further contribution to the literature, which has generally focused on returns upon announcement of addition to the index.

We further hypothesise that increased market scrutiny and information availability upon addition to the S&P 500 will aid investors in identifying market inefficiencies, in which case the degree of mispricing should, on average, decrease following addition to the index.

3 Methodology

The objective of this study is to investigate whether index premia are demand-driven by investigating whether mispricing explains a stock's returns following addition to the S&P 500. This study comprises three distinct applications of empirical techniques: the *primary regressions*, an *event study* and the *auxiliary regressions*. The research question is investigated via the primary regressions of addition premia and post-addition returns on mispricing prior to the period over which abnormal returns are estimated. The event study estimates the stock price impact of addition to the S&P 500 and post-addition returns: the dependent variables in the primary regressions. The auxiliary regressions estimate a stock's mispricing: the explanatory variables of interest in the primary regressions.

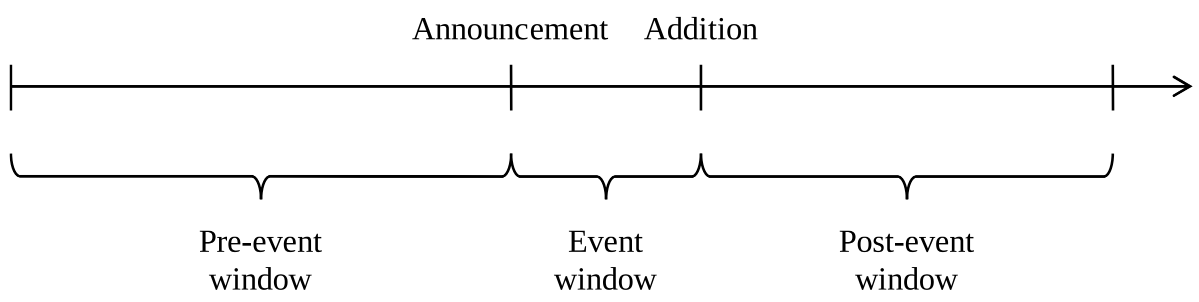
3.1 Event Study

The event study estimates stock price reactions to addition to the S&P 500 and post-addition returns: the dependent variables in the primary regressions. Compustat – Capital IQ provides constituent data for the S&P 500 including effective date of addition and deletion, dating back to 1964, however only additions from 1990 to January 2020 are considered in the event study.

3.1.1 Event Timeline

Constituent changes take effect from the effective addition day. The S&P Index Committee announces constituent changes at 5:15pm Eastern Standard Time (after close of trade for the U.S. markets) no more than five trading days prior to the effective addition day. Index-tracking investors re-balance their portfolios as of the effective addition day in order to minimise tracking error (Chen, Noronha & Singal, 2004). Figure 1 illustrates the timeline of the event study.

Figure 1: Event timeline



3.1.1.1 Event Window

Abnormal returns are estimated over a range of windows including the *event window* as well as *pre-* and *post-event windows*. The event window is defined as the period from announcement to addition. The pre-event window begins 60 days prior to the event window to capture the effect of any anticipatory buying or information leakage prior to announcement. The post-event window extends 50 days after addition to accommodate investigation into the persistence of addition premia.

3.1.1.2 Estimation Window

Models of stock returns used to estimate normal returns over the event and post-event windows are estimated over the 126 trading days (six months) preceding the event window so as not to capture any addition effect. For the pre-event windows, the estimation window is the preceding 126 trading days.

3.1.2 Normal Returns

Normal (log) returns are estimated using the market model, adjusted market model, Fama and French's (1993) 3-factor model and Carhart's (1997) 4-factor model.

3.1.2.1 Market Model

Under the market model, expected normal returns are given by equation (1).

$$E\left[R_{i,t}^{*}|\Omega_{i,t}\right] = R_{f,t}^{*} + \hat{\beta}_{i}\left(R_{m,t}^{*} - R_{f,t}^{*}\right) \tag{1}$$

$R_{i,t}^{*}$ is the return on stock i at time t, * indicates that t is within the window over which abnormal returns are being estimated, $\Omega_{i,t}$ represents conditioning information for stock i at time t, $R_{f,t}^{*}$ is the risk-free interest rate at time t, $R_{m,t}^{*}$ is the market return at time t and $\hat{\beta}_{i}$ is estimated by applying ordinary least squares to equation (2) over the estimation window.

$$\left(R_{i,t}-R_{f,t}\right) = \alpha_{i} + \beta_{i}\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{i,t} \tag{2}$$

3.1.2.2 Adjusted Market Model

Petajisto (2011) uses the adjusted market model to estimate normal returns, arguing that added stocks tend to perform well prior to addition, introducing bias in estimating α_{i}. The adjusted market model is a special case of the market model that avoids this issue by setting α_{i} equal to zero. It sets β_{i} equal to one so that the expected normal return is equal to the market return during the event window as illustrated in equation (3).

$$E\left[R_{i,t}^{*}|\Omega_{i,t}\right] = R_{m,t}^{*} \tag{3}$$

3.1.2.3 Three-Factor Model

Under Fama and French's (1993) three-factor model, expected normal returns are given by equation (4).

$$E\left[R_{i,t}^{*}|\Omega_{it}\right] = R_{f,t}^{*} + \hat{\beta}_{i,mkt}\left(R_{mkt,t}^{*} - R_{f,t}^{*}\right) + \hat{\beta}_{i,SMB}R_{SMB,t}^{*} + \hat{\beta}_{i,HML}R_{HML,t}^{*} \tag{4}$$

$R_{mkt,t}^{*}$ is the market return at time t. $R_{SMB,t}^{*}$ and $R_{HML,t}^{*}$ are the excess returns on the Fama-French factors SMB and HML at time t. $\hat{\beta}_{i,mkt}$, $\hat{\beta}_{i,SMB}$ and $\hat{\beta}_{i,HML}$ are stock i's estimated sensitivities to the three respective factors and are estimated by regressing equation (5) over the estimation window.

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,mkt}\left(R_{mkt,t} - R_{f,t}\right) + \beta_{i,SMB}R_{SMB,t} + \beta_{i,HML}R_{HML,t} + \varepsilon_{i,t} \tag{5}$$

3.1.2.4 Four-Factor Model

Under Carhart's (1997) four-factor model, expected normal returns are given by equation (6).

$$E\left[R^*_{i,t}|\Omega_{it}\right] = R^*_{f,t} + \hat{\beta}_{i,mkt}\left(R^*_{mkt,t} - R^*_{f,t}\right) + \hat{\beta}_{i,SMB}R^*_{SMB,t} + \hat{\beta}_{i,HML}R^*_{HML,t} + \hat{\beta}_{i,UMD}R^*_{UMD,t} \tag{6}$$

The four-factor model is effectively the three-factor model plus an additional factor capturing momentum (UMD). $\hat{\beta}_{i,mkt}$, $\hat{\beta}_{i,SMB}$, $\hat{\beta}_{i,HML}$ and $\hat{\beta}_{i,UMD}$ are stock i's estimated sensitivities to the four respective factors and are estimated by regressing equation (7) over the estimation window.

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,mkt}\left(R_{mkt,t} - R_{f,t}\right) + \beta_{i,SMB}R_{SMB,t} + \beta_{i,HML}R_{HML,t} + \beta_{i,UMD}R_{UMD,t} + \varepsilon_{i,t} \tag{7}$$

3.1.2.5 Data

Data for the factors and the risk-free interest rate are obtained from Wharton Research Data Services' (WRDS) Fama-French Portfolio and Factors Database. Excess market return is calculated as the value-weighted return on all stocks listed on the NYSE, AMEX and NASDAQ, minus the one-month Treasury bill rate. Stock price data are obtained from Bloomberg. Additions for which stock price data are not available over the entire estimation period are excluded, thereby excluding additions under certain circumstances including initial public offerings (IPOs) and spin-offs.

3.1.3 Abnormal Returns

Abnormal returns are the difference between a stock's actual (log) return and expected normal (log) return, calculated according to equation (8).

$$\varepsilon^*_{it} = R^*_{it} - E\left[R^*_{i,t}|\Omega_{it}\right] \tag{8}$$

Summing abnormal returns over time provides cumulative abnormal returns (CAR). Stock price data for calculating actual returns are obtained from Bloomberg. Cumulative abnormal returns are averaged over all (N) additions according to equation (9).

$$\overline{CAR} = \frac{1}{N}\sum_{i=1}^{N} CAR_i \tag{9}$$

The variance and standard error ($\bar{\sigma}$) of the average CAR are estimated using the cross-sectional approach given by equation (10).

$$var[\overline{CAR}] = \bar{\sigma}^2 = \frac{1}{N} var[CAR_i] \tag{10}$$

The cross-sectional approach assumes no clustering and that abnormal returns are not correlated across observations.

3.2 Mispricing

The auxiliary regressions provide a measure of a stock's mispricing: the explanatory variables of interest in the primary regressions. The approach follows that outlined by Bernström (2014). For each addition, a valuation multiple is regressed on appropriate conditioning variables for an added stock and its industry peers, providing an estimate of the stock's valuation conditioned on its fundamental financials. The difference between a stock's actual and fundamental multiples (the stock's residual from the regression) provides a measure of mispricing. This is adjusted by dividing by the average multiple for the industry to provide the measure of mispricing used as the explanatory variables in the primary regressions. The primary regressions are performed using index premia and post-addition returns as dependent variables. Mispricing is estimated as of the day prior to announcement and the effective addition day for the regressions of index premia and post-addition returns respectively.

3.2.1 Peers

While Bernström (2014) argues for the paramount importance of conditioning variables, he expresses a preference for a combination of both appropriate conditioning variables and relevant peers. An added firm's peers are those constituents of the S&P 400 that operate in the added firm's industry according to the Bloomberg Industry Classification Standard (BICS) as of the effective addition day. Constituent data for the S&P 400 are available on Bloomberg from 1995.

3.2.2 Valuation Multiple

Bernström (2014) argues that because multiples are tools used to identify value, value is driven by free cash flow and choice of multiple does not affect underlying cash flows, that any multiple used in conjunction with appropriate conditioning variables and peers should result in the same valuation as a correctly-applied present value model. This study considers primarily market-to-book value of equity (M/B) due to its prevalence in the literature. For robustness, the analysis is repeated using price-to-earnings (P/E). P/E is commonly used in industry (Pinto, Robinson & Stowe, 2019) and is therefore used in academic literature to capture markets' beliefs. P/E can be unreliable when earnings are very small or negative. Any outliers in the auxiliary regressions will impact the residual for the added firm and therefore the measure of mispricing used as the explanatory variable in the primary regression. P/E and M/B are therefore winsorised at 5% and 95%. These levels are justified by the relatively low number of peers. Furthermore, firms with negative earnings are excluded from the auxiliary regressions and additions with negative earnings are excluded from the primary regressions, which could potentially introduce selection bias. Data for M/B and P/E are obtained from Bloomberg.

3.2.3 Conditioning Variables

The use of appropriate conditioning variables that explain cross-sectional variation in multiples within industries is necessary to obtain an accurate estimate of fundamental value. Bernström (2014) recommends conditioning M/B on expected return on equity (ROE) and a measure of risk such as a firm's weighted average cost of capital (WACC). A firm with higher expected ROE will trade at a higher multiple while a firm with a higher WACC will trade at a lower multiple (all other things equal). This approach is consistent with the RIM (Ohlson, 1991, 1995), which conditions on ROE and cost of equity. As a proxy for expected ROE, this study uses actual ROE over the previous 12 months, obtained from Bloomberg. This value is normalised for one-time charges, providing a cleaner proxy for expected ROE. WACC depends on a firm's cost of equity, cost of debt, capital structure and tax rate. All companies operate in the U.S. and are subject to approximately the same statutory tax rate. Since cost of equity and cost of debt depend on systematic risk and capital structure, this study conditions the valuation multiple on systematic risk (beta) and leverage. Beta is estimated using daily observations over five years and leverage is calculated according to equation (11). Data for beta and leverage are obtained from Bloomberg.

$$Leverage = \frac{Short\text{– }term\ debt + Long\text{– }term\ debt}{Market\ capitalisation} \tag{11}$$

Bernström (2014) recommends a similar approach to conditioning P/E, replacing expected ROE with expected earnings per share (EPS) growth. A company with higher expected EPS growth should trade at a higher multiple. Actual EPS growth, calculated as the percent change in quarterly EPS from the previous year's corresponding quarter, obtained from Bloomberg, is used as a proxy for expected EPS growth.

All conditioning variables are winsorised at 5% and 95%. These levels are justified by the relatively low number of peers.

3.2.4 Calculating Mispricing

For each addition, the valuation multiples (M/B and P/E) are regressed on their respective conditioning variables for the added firm and its peers according to equations (12) and (13).

$$M/B_i = \alpha + \beta_1 ROE_i + \beta_2 Beta_i + \beta_3 Leverage_i + \varepsilon_i \tag{12}$$

$$P/E_i = \alpha + \beta_1 EPS\ growth_i + \beta_2 Beta_i + \beta_3 Leverage_i + \varepsilon_i \tag{13}$$

The added stock's (denoted by a) fundamental valuation multiple is that justified by current financial fundamentals, given by its fitted value from the auxiliary regression as presented in equations (14) and (15).

$$\widehat{M/B}_a = \hat{\alpha} + \hat{\beta}_1 ROE_a + \hat{\beta}_2 Beta_a + \hat{\beta}_3 Leverage_a \tag{14}$$

$$\widehat{P/E}_a = \hat{\alpha} + \hat{\beta}_1 EPS\ growth_a + \hat{\beta}_2 Beta_a + \hat{\beta}_3 Leverage_a \tag{15}$$

The difference between a stock's actual and fundamental valuation multiples (its residual) provides a measure of its mispricing. This is adjusted by dividing by the industry average multiple to account for variation between industries and across time, as illustrated in equations (16) and (17).

$$Mispricing = \frac{M/B_a - \widehat{M/B}_a}{M/B_{Industry}} = \frac{\hat{\varepsilon}_a}{M/B_{Industry}} \tag{16}$$

$$Mispricing = \frac{P/E_a - \widehat{P/E}_a}{P/E_{Industry}} = \frac{\hat{\varepsilon}_a}{P/E_{Industry}} \quad (17)$$

3.3 Primary Regressions

The primary research question is investigated via regressions of addition premia and post-addition returns on mispricing where the dependent and explanatory variables are provided by the event study and auxiliary regressions respectively. The primary regressions are performed using all four models used to estimate expected normal returns as well as mispricing estimated using M/B and P/E.

3.3.1 Regression of Addition Premia on Mispricing

The regressions of addition premia on mispricing are given by equation (18).

$$Addition\ premium_i = \alpha + \beta_1 Mispricing_i + \boldsymbol{\beta_{Control}} \boldsymbol{Control_i} + \varepsilon_i \quad (18)$$

While Compustat – Capital IQ provides constituent data for the S&P 500 dating back to 1964, Bloomberg only has dollar trading volume data (used to calculate Amihud’s (2002) illiquidity) dating back to 1996 so the sample for this regression includes only additions from 1996 to January 2020 for which all data are available.

3.3.1.1 Control Variables

The regression of index premia includes control variables including the natural logarithm of market capitalisation, idiosyncratic risk, illiquidity, analyst coverage and supply elasticity. Petajisto (2011) finds that index premia are negatively related to size and positively related to idiosyncratic risk while Schnitzler (2018) finds that index premia are higher for firms that are not members of the S&P 400 Index prior to addition to the S&P 500. Amihud’s (2002) illiquidity measure is included to control for liquidity effects. We also include the number of analyst recommendations as a proxy for investor awareness to control for the effect of increased investor awareness. Dummy variables for industry and year are also included to capture fixed industry and time effects. Market capitalisation data are obtained from Bloomberg.

3.3.1.1.1 Idiosyncratic Risk

Petajisto (2011) finds that addition premia are positively related to idiosyncratic risk, which is defined as the variation in a stock’s returns not explained by the model used to estimate normal returns, measured by the volatility of the residuals. Thus, each of the models used to estimate normal returns result in a different measure of idiosyncratic risk. Estimation of idiosyncratic risk uses the same estimation window as that used to estimate expected normal returns.

3.3.1.1.2 Amihud's Illiquidity

Amihud's (2002) illiquidity is included to control for any liquidity effect. It is calculated as the average daily ratio of absolute return to dollar trading volume (in millions of dollars) over the estimation window, as presented in Equation (19).

$$Illiquidity = \frac{1}{T}\sum_{t=1}^{T}\frac{|P_t - P_{t-1}|}{Dollar\ trading\ volume_t} \tag{19}$$

All data for calculating illiquidity are obtained from Bloomberg, which only has dollar trading volume data dating back to 1996. According to the liquidity hypothesis, the less liquid a stock is prior to addition, the greater the improvement in liquidity upon addition, resulting in a larger addition premium.

3.3.1.1.3 Analyst Coverage

Analyst coverage is the number of analyst recommendations obtained from Bloomberg, which is a proxy for investor awareness. According to the investor awareness hypothesis, more sparsely covered stocks will experience a greater improvement in investor awareness and a larger premium.

3.3.1.1.4 Supply Elasticity

Supply elasticity takes a value of one if the added stock was previously a constituent of the S&P 400 and zero otherwise. Constituent data for the S&P 400 are obtained from Bloomberg and are available from 1995.

3.3.2 Regression of Post-addition returns on Mispricing

Since the timing of any price correction depends on the market's ability to identify market inefficiencies, we investigate the relationships between mispricing and both addition premia and post-addition returns. Post-addition returns over the 10 days following the event window are regressed on mispricing according to equation (20).

$$Post\text{–}\ addition\ CAR_i = \alpha + \beta_1 Mispricing_i + \boldsymbol{\beta_{Control}Control_i} + \varepsilon_i \tag{20}$$

The estimation window used to estimate expected normal returns is the same as that used in the regression of addition premia so as not to capture the stock price effect of addition to the index.

The demand shock due to the actions of index-tracking investors should peak on the effective addition day, in which case the actions of rational investors and any resulting price correction may reasonably be expected not to occur until after the effective addition day, with the resulting price correction related to mispricing as of the addition day. Mispricing is therefore estimated in the same manner as before except that the auxiliary regressions are performed using data from the last day of the event window to capture the effect of addition to the index. Since the new measure of mispricing incorporates the addition premium, those variables with a theoretical basis for explaining index premia but not returns in general (illiquidity, supply elasticity and analyst coverage) are excluded from the regression to avoid issues of multicollinearity. Size and idiosyncratic risk are retained due to theoretical predictions and empirical findings of their abilities to explain returns (Fama & French, 1992; Merton, 1987). Data for size and idiosyncratic risk are from the day prior to the event window as before, to ensure that they do not capture any addition effect. The industry and year dummy variables are also included.

Since Bloomberg only has constituent data for the S&P 400 (used to identify peers for the auxiliary regressions) dating back to 1995, the sample for this regression includes only additions from 1995 to January 2020 for which all data are available.

Table 10 in Appendix A. presents the results of diagnostic tests of the specification of equation (20), including tests for multicollinearity and linearity.

4 Results and Discussion

4.1 Event Study

4.1.1 Addition Premia

Table 1 presents a summary of addition premia for the full sample of additions from 1990 to January 2020.

Table 1: Abnormal returns (CAR) over the event window (from announcement to the effective addition day) for each method employed to estimate normal returns, by year and specific periods of interest.

	Adjusted Market Model			Market Model			3-Factor Model			4-Factor Model		
Year(s)	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat
1990	10	-0.35%	-0.256	10	0.07%	0.051	10	0.83%	0.489	10	0.56%	0.307
1991	6	-1.12%	-0.466	6	-0.72%	-0.311	6	-1.39%	-0.616	6	-1.28%	-0.618
1992	9	4.76%	2.072	8	3.39%	1.282	8	3.35%	1.296	8	3.67%	1.355
1993	11	1.14%	0.656	11	1.03%	0.625	11	0.41%	0.249	11	0.29%	0.178
1994	14	2.21%	2.029	11	2.00%	1.495	11	2.07%	1.539	11	2.05%	1.500
1995	23	4.77%	4.559	21	4.35%	4.172	21	4.11%	3.976	21	4.10%	3.957
1996	18	5.30%	4.182	16	4.09%	2.966	16	4.54%	3.375	16	4.62%	3.489
1997	28	7.86%	6.076	27	8.47%	6.976	27	7.58%	6.173	27	7.34%	5.989
1998	35	7.19%	5.300	33	7.95%	5.944	33	8.52%	6.193	33	8.25%	6.076
1999	41	5.88%	4.418	39	6.17%	4.479	39	5.36%	4.629	39	5.20%	4.802
2000	52	5.65%	3.649	48	5.70%	3.606	48	6.04%	3.894	48	6.22%	3.836
2001	29	2.57%	1.584	26	2.64%	1.512	26	2.91%	1.379	26	2.89%	1.468
2002	23	2.84%	2.543	22	2.53%	2.939	22	2.16%	2.462	22	1.86%	2.015
2003	9	0.53%	0.555	8	1.36%	1.282	8	1.48%	1.145	8	1.82%	1.470
2004	18	3.18%	4.538	16	3.63%	5.571	16	3.37%	4.608	16	3.47%	4.831
2005	16	3.17%	2.733	15	3.64%	4.459	15	3.08%	3.931	15	3.13%	3.599
2006	32	4.09%	4.567	27	3.79%	3.815	27	3.56%	3.770	27	3.34%	3.781
2007	37	1.22%	1.422	32	1.74%	2.166	32	1.54%	2.026	32	1.47%	2.114
2008	33	3.49%	2.492	30	3.64%	2.388	30	4.21%	2.986	30	3.98%	2.766
2009	28	2.15%	2.071	27	1.85%	1.913	27	1.82%	1.742	27	1.74%	1.759
2010	15	-0.85%	-0.777	14	-0.29%	-0.224	14	-0.57%	-0.440	14	-0.57%	-0.461
2011	18	1.67%	1.092	15	0.51%	0.353	15	0.33%	0.221	15	-0.55%	-0.314
2012	18	-0.49%	-0.456	14	0.34%	0.337	14	0.33%	0.360	14	0.30%	0.315
2013	17	0.18%	0.166	13	0.78%	0.673	13	0.52%	0.431	13	1.19%	0.829
2014	14	-0.54%	-0.372	14	-0.65%	-0.440	14	-0.24%	-0.151	14	0.03%	0.019
2015	29	1.63%	1.652	21	2.04%	2.737	21	2.14%	2.839	21	2.32%	2.898
2016	35	1.21%	1.263	28	0.72%	0.860	28	0.73%	0.860	28	0.52%	0.600
2017	34	-0.70%	-0.537	26	-1.41%	-0.924	26	-1.58%	-1.004	26	-1.60%	-1.054
2018	25	-1.41%	-1.125	23	-1.43%	-1.174	23	-1.61%	-1.398	23	-1.61%	-1.372
2019	22	1.55%	1.455	20	1.64%	1.582	20	1.60%	1.640	20	1.52%	1.510
2020	1	8.91%		1	9.35%		1	5.48%		1	5.33%	
1990-2020	700	2.81%	10.534	622	2.94%	10.603	622	2.87%	10.370	622	2.81%	10.172
1996-2020	627	2.85%	9.852	555	3.03%	10.078	555	2.96%	9.893	555	2.89%	9.706
1990-2010	487	3.86%	11.686	447	3.98%	11.730	447	3.91%	11.554	447	3.83%	11.423
2010-2020	228	0.34%	0.890	189	0.24%	0.619	189	0.16%	0.409	189	0.13%	0.321

Abnormal returns are significant and positive for the full sample (1990-2020) as well as the final sample used in the primary regression of addition premia on mispricing (1996-2020), confirming the existence of an index premium and the findings of Shleifer (1986), Harris and Gurel (1986), Chen, Noronha and Singal (2004) and Petajisto (2011) among others. Upon visual inspection, addition premia appear to be related to general market conditions, experiencing declines around market events including the recession of the early 1990s, the burst of the dot-com bubble and the global financial crisis. While the decline in magnitude and significance from 2009/2010 might be attributed to the global financial crisis, the analysis is extended to estimate abnormal returns over the pre-event windows to investigate the possibility of increased information leakage or anticipatory buying due to the market's increased awareness of the addition premium.

4.1.2 Pre-Event Returns

Abnormal returns over the pre-event windows are presented in Table 2.

Table 2: Abnormal returns (CAR) over the event window (from announcement to the effective addition day) and pre-event windows (from 20 days prior to the event window). The estimation period for each window is the preceding 126 trading days (six months).

		Adjusted Market Model			**Market Model**			**3-Factor Model**			**4-Factor Model**		
Window	Sample	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat
-20 to -11	1990 - 2010	455	0.87%	2.348	444	0.66%	1.836	444	0.72%	2.011	444	0.47%	1.313
	2010 - 2020	192	0.74%	1.915	190	0.53%	1.394	190	0.36%	0.980	190	0.36%	0.993
	1990 - 2020	633	0.86%	2.954	620	0.63%	2.254	620	0.63%	2.265	620	0.47%	1.694
-10 to -1	1990 - 2010	458	1.24%	2.851	446	1.06%	2.480	446	1.04%	2.492	446	0.88%	2.086
	2010 - 2020	192	0.81%	2.523	190	0.72%	2.232	190	0.68%	2.193	190	0.49%	1.591
	1990 - 2020	636	1.09%	3.344	622	0.95%	2.970	622	0.93%	2.970	622	0.75%	2.391
Event	1990 - 2010	487	3.86%	11.686	447	3.98%	11.730	447	3.91%	11.554	447	3.83%	11.423
window	2010 - 2020	228	0.34%	0.890	189	0.24%	0.619	189	0.16%	0.409	189	0.13%	0.321
	1990 - 2020	700	2.81%	10.534	622	2.94%	10.603	622	2.87%	10.370	622	2.81%	10.172

From 2010, the small magnitude of abnormal returns over the ten days prior to the event window (-10 to -1) and the lack of significance for the previous ten days (-20 to -11) are not consistent with information leakage or anticipation. Indeed, the decline in abnormal returns in 2010 persists when abnormal returns are estimated over the event window and the previous ten days as illustrated in Table 11 in Appendix B.

4.1.3 Post-addition returns

The analysis is extended beyond the event window to investigate the persistence of addition premia as presented in Table 3.

Table 3: Abnormal returns (CAR) for the event window (from announcement to the effective addition day) and post-event windows (until 50 days after the event window). The estimation period for all windows is the 126 trading days (six months) preceding the event window.[1]

		Adjusted Market Model			**Market Model**			**3-Factor Model**			**4-Factor Model**		
Window	Sample	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat
Event	1990 - 2020	700	2.81%	10.534	622	2.94%	10.603	622	2.87%	10.370	622	2.81%	10.172
window	1996 - 2020	627	2.85%	9.852	555	3.03%	10.078	555	2.96%	9.893	555	2.89%	9.706
+1 to +10	1990 - 2020	713	-1.65%	-5.834	621	-1.31%	-4.678	621	-1.45%	-5.116	621	-1.41%	-4.852
	1995 - 2020	663	-1.77%	-6.020	575	-1.49%	-5.088	575	-1.67%	-5.683	575	-1.64%	-5.437
+11 to +20	1990 - 2020	717	-0.87%	-3.338	621	-1.12%	-4.179	621	-1.49%	-5.503	621	-1.47%	-5.401
	1995 - 2020	663	-0.98%	-3.640	575	-1.24%	-4.530	575	-1.65%	-6.015	575	-1.65%	-5.951
+21 to +30	1990 - 2020	717	-0.89%	-2.857	621	-1.07%	-3.391	621	-1.30%	-3.827	621	-1.28%	-3.911
	1995 - 2020	663	-0.89%	-2.701	575	-1.05%	-3.172	575	-1.28%	-3.580	575	-1.26%	-3.660
+31 to +40	1990 - 2020	715	-1.29%	-2.430	619	-1.39%	-2.567	619	-1.76%	-3.248	619	-1.96%	-3.368
	1995 - 2020	661	-1.29%	-2.290	573	-1.38%	-2.401	573	-1.77%	-3.090	573	-1.96%	-3.171
+41 to +50	1990 - 2020	712	-0.79%	-2.710	616	-0.95%	-3.140	616	-0.82%	-2.637	616	-0.94%	-3.046
	1995 - 2020	658	-0.80%	-2.581	570	-0.98%	-3.058	570	-0.82%	-2.490	570	-0.95%	-2.908

Abnormal returns over the post-event windows are significantly negative for the full sample (1990-2020) as well as the final sample used in the primary regression of post-addition returns on mispricing (1995-2020), supporting previous findings that addition premia experience at least a degree of reversal (Beneish & Whaley, 1996; Lynch & Mendenhall, 1997; Petajisto, 2011; Ravi & Hong, 2015). Table 4 shows that addition premia are temporary, undergoing a full reversal within the 20 days following addition.

1 Results are reported for 1996-2020 for the event window but 1995-2020 for the post event windows because the sample used in the primary regression of addition premia on mispricing is smaller due to a lack of availability of a control variable (illiquidity) that is not included in the primary regression of post-addition returns on mispricing.

Table 4: Abnormal returns (CAR) over the event window (from announcement to the effective addition day) and the following 20 days for each of the methods employed to estimate normal returns, by year and specific periods of interest.

	Adjusted Market Model			**Market Model**			**3-Factor Model**			**4-Factor Model**		
Year(s)	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat
1990	11	-0.61%	-0.232	10	0.82%	0.265	10	3.15%	0.693	10	3.04%	0.681
1991	8	0.01%	0.002	6	2.05%	0.387	6	0.62%	0.109	6	1.07%	0.184
1992	9	0.90%	0.342	8	0.13%	0.044	8	1.48%	0.430	8	2.32%	0.640
1993	11	6.12%	2.390	11	5.50%	2.206	11	5.35%	1.962	11	5.86%	2.108
1994	15	2.13%	0.874	11	2.76%	1.270	11	3.01%	1.686	11	3.15%	1.782
1995	26	-0.15%	-0.061	21	0.00%	-0.001	21	-0.40%	-0.149	21	-0.52%	-0.187
1996	18	1.43%	0.556	16	0.01%	0.005	16	-1.08%	-0.361	16	-0.84%	-0.268
1997	29	3.90%	2.353	27	3.99%	2.226	27	1.95%	1.023	27	1.59%	0.855
1998	36	1.13%	0.553	33	1.10%	0.492	33	1.91%	0.790	33	2.63%	1.038
1999	42	0.50%	0.188	39	1.21%	0.453	39	-1.34%	-0.546	39	-1.72%	-0.733
2000	52	3.16%	1.248	48	3.83%	1.575	48	1.67%	0.643	48	0.90%	0.345
2001	29	0.54%	0.190	26	1.61%	0.655	26	0.31%	0.099	26	2.93%	1.053
2002	23	-0.71%	-0.309	22	-0.91%	-0.392	22	-1.46%	-0.681	22	-1.57%	-0.749
2003	9	-1.86%	-1.224	8	-0.03%	-0.040	8	1.48%	1.368	8	1.59%	1.449
2004	18	-0.27%	-0.202	16	0.46%	0.324	16	0.21%	0.151	16	0.33%	0.239
2005	16	2.99%	1.889	15	2.76%	1.731	15	2.03%	1.312	15	1.36%	0.784
2006	32	-0.32%	-0.189	27	-0.34%	-0.215	27	-1.37%	-0.840	27	-0.32%	-0.188
2007	37	0.06%	0.032	32	1.90%	1.160	32	1.09%	0.536	32	0.61%	0.318
2008	34	-6.04%	-2.605	30	-6.37%	-2.360	30	-5.01%	-2.163	30	-6.89%	-2.943
2009	28	-1.63%	-0.795	27	-2.76%	-1.417	27	-3.18%	-1.453	27	-3.88%	-1.368
2010	15	-2.02%	-1.059	14	-2.64%	-1.119	14	-2.72%	-1.167	14	-2.34%	-0.970
2011	18	-1.07%	-0.599	15	-1.38%	-0.751	15	-1.64%	-0.867	15	-0.89%	-0.473
2012	18	-0.45%	-0.179	14	0.14%	0.053	14	0.80%	0.301	14	0.81%	0.289
2013	17	-1.23%	-0.655	13	-0.66%	-0.291	13	-1.31%	-0.544	13	-1.59%	-0.571
2014	14	0.53%	0.195	14	0.53%	0.188	14	1.11%	0.395	14	1.23%	0.439
2015	30	1.74%	1.172	21	2.18%	1.235	21	1.84%	1.169	21	2.40%	1.658
2016	37	0.95%	0.640	28	0.90%	0.635	28	0.67%	0.561	28	0.68%	0.526
2017	34	-0.70%	-0.361	26	-1.96%	-0.941	26	-2.26%	-1.106	26	-2.28%	-1.155
2018	26	-1.41%	-0.769	23	-1.43%	-0.777	23	-2.31%	-1.315	23	-2.43%	-1.283
2019	25	-0.73%	-0.328	20	1.93%	1.651	20	1.67%	1.556	20	1.44%	1.336
2020	1	4.76%		1	6.47%		1	-5.24%		1	-4.07%	
1990-2020	718	0.24%	0.548	622	0.51%	1.149	622	-0.06%	-0.136	622	-0.07%	-0.153
1995-2020	664	0.11%	0.230	576	0.36%	0.763	576	-0.31%	-0.644	576	-0.35%	-0.709
1990-2010	498	0.40%	0.714	447	0.69%	1.208	447	0.01%	0.009	447	-0.04%	-0.066
2010-2020	235	-0.25%	-0.404	189	-0.13%	-0.214	189	-0.42%	-0.699	189	-0.31%	-0.513

The complete reversal in addition premia is consistent with the downward-sloping short-run demand curve hypothesis and the findings of Harris and Gurel (1986), Vespro (2006) and Patel and Welch (2017). It is not consistent with Shleifer's (1986) downward-sloping long-run demand curve hypothesis or the information-based theories of the index premium, which predict a persistent effect. Furthermore, it is consistent with our hypothesis that the addition effect is demand-driven, leading to overpricing and subsequent selling by rational investors.

4.2 Mispricing

A summary of the results of the auxiliary regressions are presented in Table 5.

Table 5: Summary statistics for the auxiliary regressions, including resulting measures of mispricing and regression statistics. The regressions are performed and mispricing estimated, using M/B and P/E, one day prior to the event window and one day prior to the post-event window.

	Regressions				Mispricing						
		Min.	Max.	Mean				Mean			
	R^2	Obs.	Obs.	Obs.	Obs.	Min.	Max.	Mean	Test stat.	Median	SD.
M/B											
1d prior to event window	0.510	7	101	44.986	512	-2.169	4.363	0.230	7.858	0.107	0.663
1d prior to post-event window	0.508	7	102	45.053	512	-2.230	4.466	0.236	8.041	0.121	0.663
P/E											
1d prior to event window	0.226	6	98	43.749	538	-1.828	4.682	0.175	5.901	0.035	0.689
1d prior to post-event window	0.222	6	97	43.742	538	-1.747	4.052	0.163	5.582	0.015	0.678

The quality of the measure of mispricing depends on the ability of the conditioning variables to explain cross-sectional variation in the valuation multiple within industries. The average R-squared for the auxiliary regressions for M/B are significantly higher than for P/E, suggesting that using M/B provides a more accurate estimate of mispricing. The auxiliary regressions are performed with a minimum number of observations of 7 (6) and an average of 45 (44) for M/B (P/E), in line with Bernström's (2014) guidance that 5 to 10 relevant peers should be sufficient for an accurate valuation. While this is relatively low, it is validated by the purpose of the auxiliary regressions being to capture the market's beliefs regarding added stocks' fundamental valuations.

Average mispricing is significantly positive prior to addition with test statistics of 7.86 and 5.90 for M/B and P/E respectively. This is potentially consistent with Petajisto's (2011) observation that firms are added to the index when they have exhibited strong performance. Assuming that addition premia are demand-driven, valuations are expected to increase following addition, which is indeed what is observed, with average mispricing estimated using M/B increasing to 0.24 with a test statistic of 8.04. Conversely, average mispricing estimated using P/E decreases to 0.16 with a test statistic of 5.58, however, given P/E's inferior performance in the auxiliary regressions according to the average R-squared, more weight is given to M/B.

The square of mispricing provides a measure of the absolute degree of mispricing for each added stock. The average of this measure estimated using M/B decreases 11% (from 0.50 to 0.44) from prior to announcement to 30 trading days after the effective addition day. This is consistent with the hypothesis that increased market scrutiny and information availability upon addition to the S&P 500 aids investors in identifying and eliminating market inefficiencies.

4.3 Primary Regressions

The timing of any price correction depends on the market's ability to identify market inefficiencies. As such, we investigate the relationships between mispricing and addition premia and between mispricing and post-addition returns. We primarily consider mispricing estimated using M/B on the basis that its conditioning variables are better able to explain cross-sectional variation compared to P/E, providing a more accurate estimate of mispricing.

4.3.1 Addition Premia

Table 6 presents summary statistics for the data used in the regression of addition premia on mispricing estimated using M/B.

Table 6: Summary statistics for data used in the regression of addition premia (estimated over the event window) on mispricing. Mispricing is estimated using data from immediately prior to the event window using M/B.

	Obs.	Min.	Max.	Mean	Median	SD.
Dependent Variable						
CAR						
Adjusted market model	484	-0.329	0.383	0.029	0.021	0.072
Market model	484	-0.267	0.361	0.029	0.021	0.071
3-factor model	484	-0.236	0.339	0.029	0.021	0.069
4-factor model	484	-0.233	0.372	0.028	0.021	0.069
Explanatory Variable of Interest						
Mispricing	484	-2.169	4.363	0.236	0.116	0.649
Control Variables						
Size	484	6.430	11.856	9.051	8.989	0.658
Idiosyncratic risk						
Adjusted market model	484	0.007	0.090	0.022	0.019	0.012
Market model	484	0.006	0.090	0.021	0.018	0.012
3-factor model	484	0.006	0.090	0.021	0.018	0.011
4-factor model	484	0.006	0.090	0.020	0.018	0.011
Illiquidity	484	0.000	0.232	0.003	0.001	0.013
Analyst coverage	484	0	41	12.816	12	8.460
Supply elasticity	484	0	1	0.622	1	0.485

The distribution of abnormal returns and idiosyncratic risk are robust to the method of estimating normal returns. Average abnormal returns are positive, consistent with the findings of the event study, the existence of an index premium and previous research (Shleifer, 1986; Harris & Gurel, 1986; Chen, Noronha & Singal, 2004; Petajisto, 2011). As reported in the results of the auxiliary regressions, average mispricing prior to the event window is positive, which is potentially consistent with Petajisto's (2011) observation that added firms have experienced strong performance. Furthermore, 62% of additions were previously constituents of the S&P 400, consistent with Schnitzler's (2018) 58%.

The results of the regression of addition premia on mispricing are presented in Table 7.

*Table 7: Results of the regression of addition premia (estimated over the event window) on mispricing. Mispricing is estimated using M/B and all explanatory variables are calculated using data from immediately prior to the event window. P-values reported in parentheses, calculated using White's heteroskedasticity-consistent standard errors. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	**Market Model**	**3-Factor Model**	**4-Factor Model**
Intercept	0.099	0.090	0.050	0.051
	(0.235)	*(0.249)*	*(0.432)*	*(0.414)*
Mispricing	-0.001	-0.004	-0.009*	-0.009*
	(0.902)	*(0.392)*	*(0.063)*	*(0.079)*
Size	-0.002	0.000	0.003	0.003
	(0.845)	*(0.980)*	*(0.722)*	*(0.703)*
Idiosyncratic risk	1.181**	1.052**	1.108**	1.004*
	(0.018)	*(0.036)*	*(0.049)*	*(0.072)*
Illiquidity	0.361**	0.158	0.158	0.175
	(0.026)	*(0.199)*	*(0.193)*	*(0.165)*
Analyst coverage	0.001*	0.001**	0.001**	0.001**
	(0.096)	*(0.029)*	*(0.050)*	*(0.048)*
Supply elasticity	-0.045**	-0.048**	-0.048**	-0.047**
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.295	0.308	0.323	0.306
Observations	484	484	484	484

Mispricing is insignificant at the 5% level for all models but significant at the 10% level for two models. This lack of significance is consistent with demand-based theories, under which index premia are determined by the magnitude of the demand shock and the elasticities of supply and demand, and are therefore independent of mispricing. Furthermore, the mechanical buying by index-tracking investors is independent of mispricing. According to our hypothesis, it is the actions of rational investors that are driven by mispricing. Rational investors should not act until after the demand shock due to the mechanical buying by index-tracking investors, which occurs on the effective addition day. As such, the actions of rational investors should occur during the period following the effective addition day, consistent with our finding that index premia undergo a reversal in the period following addition and the findings of previous research (Harris & Gurel, 1986; Vespro, 2006; Patel & Welch, 2017). For this reason, in Section 4.3.2, we extend our analysis to investigate the relationship between mispricing and post-addition returns.

While previous research has found size to be negatively related to returns (Fama & French, 1992), we find it to be insignificant across all models. However, these studies typically consider returns over the longer-term and in the absence of a specific event. Indeed, our findings are consistent with those of Schnitzler (2018) and Petajisto (2011), who finds only weak significance.

Idiosyncratic risk is significantly positive at the 5% level for three of the four models and significant at the 10% level for the remaining model. This is consistent with theoretical predictions that investors are rewarded for exposure to risk (Merton, 1987) and the findings of Petajisto (2011) and Schnitzler (2018).

Illiquidity is significantly positive at the 5% level for the adjusted market model, consistent with the liquidity hypothesis (Dhillon & Johnson, 1991; Beneish & Gardner, 1995; Hegde & McDermott, 2003; Ravi & Hong, 2015) which proposes that higher illiquidity prior to addition results in greater improvement in liquidity upon addition, justifying a larger premium. However, illiquidity is insignificant for the remaining models, consistent with the findings of Schnitzler (2018). Furthermore, this is consistent with our findings from the event study, that addition premia undergo a complete reversal, since the liquidity hypothesis predicts that addition premia should be persistent.

Analyst coverage is significantly positive at the 5% level under three models and at the 10% level for the remaining model, contrary to the investor awareness hypothesis and supporting research (Chen, Noronha & Singal, 2004; Denis et al., 2003; Ravi & Hong, 2015), which proposes that the improved profile due to addition to the S&P 500 improves a firm's access to capital, increasing fundamental value.

Supply elasticity is significantly negative at the 5% level for all models, consistent with the findings of Schnitzler (2018), who hypothesises that stocks that are simultaneously added to the S&P 500 and deleted from the S&P 400 have more elastic supply, resulting in a smaller addition premium.

The results of the regression of addition premia on mispricing are robust to the choice of multiple, with the above findings confirmed by regressing addition premia on mispricing estimated using P/E (Table 12, Appendix C.). Mispricing and size are insignificant for all models at the 10% level, idiosyncratic risk is significant at 5% for two models and at 10% for the remaining two. Illiquidity is significantly positive at the 5% level under the adjusted market model and at the 10% level under the four-factor model. Analyst coverage is significantly positive at the 5% level under two models and at the 10% level under one model. Supply elasticity is significantly negative at the 5% level for all models.

4.3.2 Post-Event Window

Since we find that index premia are independent of mispricing, suggesting that any price correction due to the actions of rational investors occur after the effective addition day, we regress post-addition returns, calculated over the ten days following the event window, on mispricing. Table 8 presents summary statistics for the data used in the regression of post-addition returns on mispricing estimated using M/B.

Table 8: Summary statistics for data used in the regression of post-addition returns (estimated over the 10 days following the event window) on mispricing. Mispricing is estimated using data from the last day of the event window using M/B.

	Obs.	Min.	Max.	Mean	Median	SD.
Dependent Variable						
CAR						
Adjusted market model	501	-0.384	0.258	-0.014	-0.008	0.075
Market model	501	-0.384	0.236	-0.015	-0.008	0.070
3-factor model	501	-0.336	0.309	-0.016	-0.008	0.071
4-factor model	501	-0.341	0.318	-0.016	-0.007	0.073
Explanatory Variable of Interest						
Mispricing	501	-2.230	4.466	0.229	0.121	0.649
Control Variables						
Size	501	6.430	11.856	9.028	8.967	0.665
Idiosyncratic risk						
Adjusted market model	501	0.007	0.090	0.022	0.019	0.012
Market model	501	0.006	0.090	0.021	0.018	0.012
3-factor model	501	0.006	0.090	0.021	0.018	0.011
4-factor model	501	0.006	0.090	0.020	0.018	0.011

The distribution of abnormal returns and idiosyncratic risk are robust to the model used to estimate normal returns. Average abnormal returns are negative, consistent with our findings in the event study and previous findings (Harris & Gurel, 1986; Vespro, 2006; Patel & Welch, 2017) of a reversal of index premia following addition. Average mispricing after addition is positive, consistent with our findings that an index premium exists (and that average mispricing prior to addition is positive): an overvalued stock that experiences a price increase is expected to become more overvalued (in the absence of a change in fundamental financials). However, comparing Table 8 and Table 6, mispricing appears to decrease over the event window. Because the regressions use different control variables and because observations with missing data are excluded, the two regressions use slightly different samples. Indeed, average mispricing for the sample used in the regression of addition premia on mispricing, increases from 0.236 to 0.240 over the event window, in line with expectations. Note that this is consistent with the findings presented in Table 5, the sample for which includes additions from 1995 to January 2020 and includes all observations excluded from the regressions based on lack of data.

Table 9 presents the results of the regression of post-addition returns on mispricing.

*Table 9: Results of the regression of post-addition returns (estimated over the 10 days following the event window) on mispricing. Mispricing is estimated using M/B as of the last day of the event window. Size and idiosyncratic risk are taken from the day prior to the event window. P-values reported in parentheses, calculated using White's heteroskedasticity-consistent standard errors. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	**Market Model**	**3-Factor Model**	**4-Factor Model**
Intercept	-0.041	-0.053	-0.047	-0.054
	(0.504)	*(0.372)*	*(0.489)*	*(0.448)*
Mispricing	-0.014**	-0.014**	-0.009**	-0.014**
	(0.010)	*(0.004)*	*(0.040)*	*(0.005)*
Size	0.009	0.009	0.008	0.008
	(0.186)	*(0.156)*	*(0.253)*	*(0.276)*
Idiosyncratic risk	-0.807*	-0.331	-0.651	-0.262
	(0.074)	*(0.388)*	*(0.158)*	*(0.610)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.104	0.121	0.111	0.114
Observations	501	501	501	501

Abnormal returns are inversely related to mispricing at the 5% level for all models, consistent with our hypothesis that addition to the S&P 500 results in short-term mispricing, which undergoes a subsequent correction via the actions of rational investors. An increase of one standard deviation in mispricing corresponds to a decrease in abnormal returns of between 0.09 and 0.13 standard deviations depending on the model used to estimate normal returns. This reflects the noisy nature of returns, which are influenced by many factors and are therefore difficult to explain, especially in the absence of a specific event such as during the post-event period. This is also reflected in the low R-squared values. These results are robust to the length of the post-event window, with mispricing remaining significantly negative at the 5% level for all models when post-addition returns are estimated over 30 days (Table 13, Appendix D.).

Contrary to empirical findings (Fama & French, 1992; Merton, 1987), size and idiosyncratic risk are insignificant for all models, possibly due to the noisy nature of returns, especially over a relatively short window.

In the regression of post-addition returns on mispricing estimated using P/E, all explanatory variables are insignificant (Table 14, Appendix E.). However, when the window over which post-addition returns are estimated is extended to 30 days, mispricing is significant at the 5% level for two models and 10% for the remaining two (Table 15, Appendix F.). Given the limitations of P/E outlined in Section 3.2.2 and its inferior ability to explain cross-sectional variation compared to M/B, we are wary of placing too much emphasis on the results of the analysis based on P/E. These limitations include the challenge of winsorising a small sample. It is possible that we fail to adjust all outliers, which will affect the measure of mispricing and therefore the results of the primary regressions. Furthermore, excluding companies with negative earnings could introduce a selection bias into the auxiliary regressions, further affecting our measure of mispricing.

5 Conclusion

The purpose of this study is to use a novel approach to estimating mispricing to investigate the positive stock price reaction to addition to the S&P 500. Building on the downward-sloping short-run demand curve hypothesis of index premia (Harris & Gurel, 1986), we hypothesise that addition to the index results in short-term mispricing, which undergoes a subsequent correction via the actions of rational investors.

We find evidence for the existence of an index premium, confirming the findings of previous research. Index premia appear to be cyclical, although they have failed to recover since the global financial crisis, representing a possible subject for further research in this area. We find that addition premia are temporary, undergoing a complete reversal within 20 days, consistent with the hypothesis that demand curves are downward-sloping in the short-term and inconsistent with the downward-sloping long-run demand curve and information-based theories of the index premium, which postulate that the effect should be persistent. These findings are also consistent with our hypothesis that, if addition premia are indeed demand-driven, determined by the magnitude of the demand shock due to addition to the index and the elasticities of supply and demand, that this represents a departure from market efficiency, which will be identified and eliminated via the actions of rational investors, resulting in a subsequent price correction.

In the primary regressions, we find that addition premia are independent of mispricing. This is consistent with our theory that the price correction does not occur until after the demand shock due to the actions of index-tracking investors, which occurs on the effective addition day. It is also consistent with the downward-sloping short-run demand curve hypothesis, under which index premia are determined by the magnitude of the demand shock and the elasticities of supply and demand and are therefore independent of mispricing. As expected, post-addition returns are found to be negatively related to mispricing, suggesting that the market behaves rationally and is able to identify and eliminate inefficiencies. We speculate that this may be due to increased market scrutiny and information availability upon addition to the index, aiding investors in identifying pricing inefficiencies, finding that absolute mispricing for added stocks decreases over the period from prior to announcement to 30 days following addition. The relationship between information availability and mispricing represents a possible subject for further research, with stocks that experience a greater improvement in information availability expected to experience a greater decrease in absolute mispricing.

We found post-addition returns over the 10 days following addition to be independent of mispricing estimated using P/E, and a negative relationship when the post-addition window is extended to 30 days. Alternative multiples and approaches may better capture the market's belief and test the robustness of our results. We discuss the trade-off between quality and quantity of peers in the auxiliary regressions. Bernström (2014) argues that the use of appropriate value drivers makes the choice of peers redundant. Such an approach would result in a larger number of observations used in the auxiliary regressions and may provide a more reliable measure of mispricing. Furthermore, the auxiliary regressions rely on measuring market expectations. Not only are broker forecasts imperfect proxies for market expectations, but due to data availability, we were forced to use actual values as proxies for forecasts and expectations. This issue is exacerbated by the relatively low frequency of constituent changes, at least partly due to the S&P 500 Index Committee's stated objective of avoiding constituent turnover (S&P Dow Jones Indices, 2020a). While we consider post-event windows of varying lengths, further research might investigate the degree of market efficiency following addition to the S&P 500 by measuring how long it takes for the actions of rational investors to eliminate inefficiencies following addition and whether this is related to mispricing, information availability or both. Finally, while many studies have investigated the effects of constituent changes for various indices, the ubiquity of the S&P 500 due to its status as a proxy for the U.S. market and the size of the U.S. market means that it is challenging to confirm these findings via investigation of alternative indices.

References

Amihud, Y. (2002). Illiquidity and Stock Returns: Cross-section and time-series effects. *Journal of Financial Markets*, vol. 5, no. 1, pp. 31-56

Amihud, Y., & Mendelson, H. (1986). Asset Pricing and the Bid-Ask Spread. *Journal of Financial Economics*, vol. 17, no. 2, pp. 223-249

Andrade, G., Mitchell, M., & Stafford, E. (2001). New Evidence and Perspectives on Mergers. *Journal of Economic Perspectives*, vol. 15, no. 2, pp. 103-120

Asquith, P., & Mullins Jr, D. W. (1986). Equity Issues and Offering Dilution. *Journal of Financial Economics*, vol. 15, pp. 61-89

Beneish, M. D., & Gardner, J. C. (1995). Information Costs and Liquidity Effects from Changes in the Dow Jones Industrial Average List. *Journal of Financial and Quantitative Analysis*, vol. 30, no. 1, pp.135-157

Beneish, M. D., & Whaley, R. E. (1996). An Anatomy of the "S&P Game": The effects of changing the game. *Journal of Finance,* vol. 51, no. 5, pp. 1909–1930

Bernström, S. (2014). Valuation: The market approach, John Wiley & Sons

Cai, J., & Houge, CFA, T. (2008). Long-Term Impact of Russell 2000 Index Rebalancing. *Financial Analysts Journal*, vol. 64 no. 4, pp. 76-91

Carhart, M. (1997). On Persistence in Mutual Fund Returns. *Journal of Finance,* vol. 52, no. 1, pp. 57–82

Chang, E. C., Luo, Y., & Ren, J. (2013). Pricing Deviation, Misvaluation Comovement, and Macroeconomic Conditions. *Journal of Banking & Finance*, vol. 37, no. 12, pp. 5285-5299

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. *The Journal of Finance*, vol. 59, no. 4, pp. 1901-1930

Cootner, P. H. (1964). The Random Character of Stock Market Prices. Cambridge: M.I.T. Press

D'Mello, R., & Shroff, P. K. (2000). Equity Undervaluation and Decisions Related to Repurchase Tender Offers: An empirical investigation. *The Journal of Finance*, vol. 55, no.5, pp. 2399-2424

De Bondt, W. F., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, vol. 40, no. 3, pp. 793-805

De Bondt, W. F., & Thaler, R. H. (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality. *The Journal of Finance*, vol. 42, no. 3, pp. 557-581

Denis, D. K., McConnell, J. J., Ovtchinnikov, A. V., & Yu, Y. (2003). S&P 500 Index Additions and Earnings Expectations. *The Journal of Finance*, vol. 58, no. 5, pp. 1821-1840

Dhillon, U., & Johnson, H. (1991). Changes in the Standard and Poor's 500 List. *The Journal of Business*, vol. 64, no. 1, pp. 75-85

Dong, M., Hirshleifer, D., & Teoh, S. H. (2012). Overvalued Equity and Financing Decisions. *The Review of Financial Studies*, vol. 25, no. 12, pp. 3645-3683

Doukas, J. A., Kim, C. F., & Pantzalis, C. (2010). Arbitrage Risk and Stock Mispricing. *Journal of Financial and Quantitative Analysis*, vol. 45, no. 4, pp. 907-934

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, vol. 47, no. 2, pp. 427-465

Fama, E. F., & French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, vol. 33, no. 1, pp. 3-56

Fernandes, M., & Mergulhão, J. (2016). Anticipatory Effects in the FTSE 100 Index Revisions. *Journal of Empirical Finance*, vol. 37, pp. 79-90

Frankel, R. M., & Lee C. (1998). Accounting Valuation, Market Expectation, and the Book-to-Market Effect. *Journal of Accounting and Economics*, vol. 25, no. 3, pp. 283–319

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. *The Journal of Finance*, vol. 41, no. 4, pp. 815-829

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*, vol. 6, no. 3, pp. 413-459

Jain, P.C. (1987). The Effect on Stock Price of Inclusion or Exclusion from the S&P 500. *Financial Analysts Journal,* vol. 43, no. 1, pp. 58–65

Kendall, M. G., & Hill, A. B. (1953). The Analysis of Economic Time-Series-Part I: Prices. *Journal of the Royal Statistical Society*. *Series A (General)*, vol. 116, no. 1, pp. 11-34

Keynes J. M. (1936). The General Theory of Employment, Interest and Money. New York: Harcourt, Brace and Co.

Lee, C. M., Myers, J., & Swaminathan, B. (1999). What is the Intrinsic Value of the Dow? *The Journal of Finance*, vol. 54, no. 5, pp. 1693-1741

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, vol. 47, pp. 13-37

Liu, J., Nissim, D., & Thomas, J. (2002). Equity Valuation Using Multiples. *Journal of Accounting Research*, vol. 40, no. 1, pp. 135-172

Lynch, A. W., & Mendenhall, R. R. (1997). New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index. *The Journal of Business*, vol. 70, no. 3, pp. 351-383

Mase, B. (2007). The Impact of Changes in the FTSE 100 Index. *Financial Review*, vol. 42, no. 3, pp. 461-484

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, vol. 42, no. 3, pp. 483-510

Mikkelson, W. H. & Partch, M. M. (1985). Stock Price Effects and Costs of Secondary Distributions. *Journal of Financial Economics*, vol. 14, no. 2, pp. 165-194

Myers, S. C., & Majluf, N. S. (1984). Corporate Financing and Investment Decisions when Firms Have Information that Investors Do not Have. *Journal of Financial Economics*, vol. 13, no. 2, pp.187-221

Ohlson, J. A. (1991). The Theory of Value and Earnings, and an Introduction to the Ball-Brown Analysis. *Contemporary Accounting Research*, vol. 8, no. 1, pp. 1-19

Ohlson, J. (1995). Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research*, vol. 11, no. 2, pp. 661–687

Patel, N., & Welch, I. (2017). Extended Stock Returns in Response to S&P 500 Index Changes. *The Review of Asset Pricing Studies*, vol. 7, no. 2, pp. 172-208

Penman, S. H., & Sougiannis, T. (1998). A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation. *Contemporary accounting research*, vol. 15, no. 3, pp. 343-383

Petajisto, A. (2011). The Index Premium and its Hidden Cost for Index Funds. *Journal of Empirical Finance*, vol. 18, no. 2, pp. 271-288

Pinto, J. E., Robinson, T. R., & Stowe, J. D. (2019). Equity Valuation: A survey of professional practice. *Review of Financial Economics*, vol. 37, no. 2, pp. 219-233

Platikanova, P. (2008). Long-Term Price Effect of S&P 500 Addition and Earnings Quality. *Financial Analysts Journal*, vol. 64, no. 5, pp. 62-76

Plenborg, T., & Pimentel, R. C. (2016). Best Practices in Applying Multiples for Valuation Purposes. *The Journal of Private Equity*, vol. 19, no. 3, pp. 55-64

Ravi, R., & Hong, Y. (2015). Information Asymmetry Around S&P 500 Index Changes. *Review of Accounting and Finance*, vol. 14, no. 2, pp. 106-127

Rhodes–Kropf, M., Robinson, D. T., & Viswanathan, S. (2005). Valuation Waves and Merger Activity: The empirical evidence. *Journal of Financial Economics*, vol. 77, no. 3, pp. 561-603

Rhodes-Kropf, M., & Viswanathan, S. (2004). Market Valuation and Merger Waves. *The Journal of Finance*, vol. 59, no. 6, pp. 2685-2718

Roberts, H. V. (1959). Stock-Market "Patterns" and Financial Analysis: Methodological suggestions. *The Journal of Finance*, vol. 14, no. 1, pp. 1-10

S&P Dow Jones Indices. (2020a). S&P U.S. Indices Methodology, Available online: https://us.spindices.com/documents/methodologies/methodology-index-math.pdf [Accessed 19 April 2020]

S&P Dow Jones Indices. (2020b). SPIVA® U.S. Scorecard, Available online: https://us.spindices.com/documents/spiva/spiva-us-year-end-2019.pdf [Accessed 21 May 2020]

Schnitzler, J. (2018). S&P 500 Inclusions and Stock Supply. *Journal of Empirical Finance*, vol. 48, pp. 341-356

Scholes, M. S. (1972). The Market for Securities: Substitution versus price pressure and the effects of information on share prices. *The Journal of Business*, vol. 45, no. 2, pp. 179-211

Sharpe, W. F. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk. The Journal of Finance, vol. 19, no. 3, pp. 425-442

Shiller, R. J. (1979). The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure. *Journal of Political Economy*, vol. 87, no. 6, pp. 1190-1219

Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review*, vol. 71, pp. 421-436

Shleifer, A. (1986). Do Demand Curves for Stocks Slope Down? *The Journal of Finance*, vol. 41, no. 3, pp. 579-590

Shleifer, A., & Vishny, R. W. (2003). Stock Market Driven Acquisitions. *Journal of Financial Economics*, vol. 70, no. 3, pp. 295-311

Summers, L. H. (1986). Does the Stock Market Rationally Reflect Fundamental Values? *The Journal of Finance*, vol. 41, no. 3, pp. 591-601

Vespro, C. (2006). Stock Price and Volume Effects Associated with Compositional Changes in European Stock Indices. *European Financial Management*, vol. 12, no. 1, pp. 103-127

Warr, R. S., Elliott, W. B., Koëter-Kant, J., & Öztekin, Ö. (2012). Equity Mispricing and Leverage Adjustment Costs. *Journal of Financial and Quantitative Analysis*, vol. 47, no. 3, pp. 589-616

Wurgler, J., & Zhuravskaya, E. (2002). Does Arbitrage Flatten Demand Curves for Stocks? *The Journal of Business*, vol. 75, no. 4, pp. 583-608

Appendices

Appendix A.

Table 10: Correlations between explanatory variables, variance inflation factors and results of the Ramsey RESET test for the model used in the regression of post-addition returns on mispricing estimated using M/B (equation (20)).

Correlation Table	Mispricing	Size
Mispricing	1.00	
Size	0.16	1.00
Idiosyncratic risk		
Adjusted market model	0.14	-0.11
Market model	0.15	-0.11
3-factor model	0.14	-0.11
4-factor model	0.14	-0.12

Variance Inflation Factor (VIF)	VIF
CAR	
Adjusted market model	1.12
Market model	1.14
3-factor model	1.12
4-factor model	1.13

Ramsey RESET Test	Adjusted Market Model	Market Model	3-Factor Model	4-Factor Model
F-value	1.12	0.15	0.27	0.01
P-value	0.33	0.86	0.77	0.99

Appendix B.

Table 11: Abnormal returns (CAR) over the event window (from announcement to the effective addition day) and the previous ten days for each method employed to estimate normal returns, by year and specific periods of interest.

	Adjusted Market Model			**Market Model**			**3-Factor Model**			**4-Factor Model**		
Year(s)	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat	Obs.	CAR	t-stat
1990	10	-2.83%	-0.933	10	-2.04%	-0.757	10	-0.98%	-0.342	10	-1.33%	-0.462
1991	6	-0.03%	-0.007	6	0.44%	0.101	6	-0.86%	-0.205	6	-0.55%	-0.131
1992	9	2.44%	0.888	8	0.53%	0.193	8	1.35%	0.557	8	1.47%	0.575
1993	11	-1.37%	-0.492	11	-1.33%	-0.458	11	-1.88%	-0.635	11	-0.76%	-0.226
1994	14	2.10%	1.468	11	2.95%	1.679	11	3.24%	1.780	11	2.81%	1.470
1995	23	5.92%	3.573	21	4.92%	2.503	21	4.70%	2.696	21	4.50%	2.479
1996	18	8.90%	5.681	16	7.20%	3.337	16	7.47%	3.266	16	7.40%	3.298
1997	28	7.06%	3.617	27	7.75%	3.929	27	6.40%	3.339	27	5.95%	2.969
1998	35	7.51%	3.800	33	8.18%	4.224	33	10.09%	5.021	33	10.15%	4.855
1999	41	10.31%	5.828	39	10.05%	5.679	39	8.79%	5.598	39	8.72%	5.713
2000	52	9.61%	4.483	48	9.38%	4.194	48	9.73%	4.944	48	9.25%	4.599
2001	29	4.68%	1.477	26	4.30%	1.404	26	3.86%	1.060	26	3.86%	1.189
2002	23	5.97%	4.211	22	5.99%	4.488	22	5.35%	3.863	22	4.74%	2.974
2003	9	0.91%	0.386	8	1.49%	0.525	8	1.76%	0.518	8	1.62%	0.499
2004	18	4.65%	5.018	16	5.26%	7.469	16	4.88%	5.051	16	4.91%	5.684
2005	16	5.11%	2.115	15	5.83%	2.528	15	4.68%	2.116	15	4.43%	1.974
2006	32	3.22%	2.708	27	2.72%	2.103	27	2.75%	2.220	27	2.78%	2.209
2007	37	1.96%	1.659	32	2.53%	2.008	32	2.24%	1.918	32	1.61%	1.435
2008	33	4.48%	2.006	30	5.12%	1.907	30	5.40%	2.248	30	4.49%	1.953
2009	28	-0.95%	-0.253	27	-1.47%	-0.395	27	-1.14%	-0.308	27	-1.39%	-0.370
2010	15	1.03%	1.057	14	0.88%	0.691	14	0.51%	0.417	14	0.55%	0.449
2011	18	2.36%	1.415	15	1.50%	0.853	15	1.41%	0.765	15	0.64%	0.316
2012	18	-0.24%	-0.183	14	0.25%	0.179	14	0.16%	0.116	14	-0.07%	-0.048
2013	17	1.92%	1.152	13	3.01%	1.522	13	2.69%	1.329	13	2.76%	1.405
2014	14	1.06%	0.560	14	0.58%	0.330	14	1.28%	0.629	14	1.49%	0.740
2015	29	2.86%	2.191	21	3.75%	2.710	21	3.46%	2.588	21	2.88%	2.264
2016	35	0.91%	0.717	28	0.58%	0.471	28	0.92%	0.741	28	0.46%	0.368
2017	34	-0.64%	-0.418	26	-1.45%	-0.779	26	-1.51%	-0.807	26	-1.63%	-0.911
2018	25	0.91%	0.651	23	1.04%	0.687	23	0.29%	0.215	23	0.17%	0.125
2019	22	0.33%	0.295	20	0.29%	0.285	20	0.63%	0.640	20	0.55%	0.509
2020	2	5.00%	0.544	2	4.61%	0.495	2	-1.02%	-0.275	2	-1.05%	-0.271
1990-2020	701	3.80%	9.409	623	3.89%	9.003	623	3.79%	8.944	623	3.55%	8.430
1996-2020	628	4.01%	9.224	556	4.16%	8.946	556	4.05%	8.853	556	3.78%	8.330
1990-2010	487	5.02%	9.456	447	5.04%	9.062	447	4.94%	9.067	447	4.71%	8.694
2010-2020	229	1.02%	2.214	190	0.96%	1.933	190	0.84%	1.720	190	0.62%	1.262

Appendix C.

*Table 12: Results of the regression of addition premia on mispricing (estimated using P/E) and control variables. Addition premia are estimated from announcement to the effective addition day. P-values in parentheses. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	Market Model	3-Factor Model	4-Factor Model
Intercept	0.082	0.087	0.080	0.081
	(0.327)	*(0.272)*	*(0.236)*	*(0.214)*
Mispricing	0.000	-0.002	-0.005	-0.004
	(0.941)	*(0.670)*	*(0.291)*	*(0.403)*
Size	-0.002	-0.001	0.001	0.000
	(0.837)	*(0.906)*	*(0.933)*	*(0.979)*
Idiosyncratic risk	1.224**	1.051**	1.045*	0.901*
	(0.011)	*(0.029)*	*(0.057)*	*(0.098)*
Illiquidity	0.360**	0.176	0.204	0.220*
	(0.029)	*(0.158)*	*(0.107)*	*(0.095)*
Analyst coverage	0.001	0.001**	0.001**	0.001*
	(0.101)	*(0.029)*	*(0.043)*	*(0.051)*
Supply elasticity	-0.045**	-0.046**	-0.047**	-0.044**
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.298	0.307	0.311	0.292
Observations	506	506	506	506

Appendix D.

*Table 13: Results of the regression of post-addition returns over the 30 days following the event window, on mispricing. Mispricing is estimated using M/B as of the last day of the event window. Size and idiosyncratic risk are taken from the day prior to the event window. P-values reported in parentheses, calculated using White's heteroskedasticity-consistent standard errors. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	**Market Model**	**3-Factor Model**	**4-Factor Model**
Intercept	0.058	0.065	0.024	0.020
	(0.538)	*(0.497)*	*(0.823)*	*(0.848)*
Mispricing	-0.023**	-0.020**	-0.020**	-0.019**
	(0.014)	*(0.021)*	*(0.028)*	*(0.047)*
Size	-0.001	-0.004	0.001	-0.001
	(0.908)	*(0.687)*	*(0.929)*	*(0.937)*
Idiosyncratic risk	-3.032**	-2.249**	-2.948**	-1.695
	(0.007)	*(0.014)*	*(0.009)*	*(0.164)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.146	0.137	0.150	0.142
Observations	501	501	501	501

Appendix E.

*Table 14: Results of the regression of post-addition returns estimated over the 10 days following the event window, on mispricing. Mispricing is estimated using P/E as of the last day of the event window. Size and idiosyncratic risk are taken from the day prior to the event window. P-values reported in parentheses, calculated using White's heteroskedasticity-consistent standard errors. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	**Market Model**	**3-Factor Model**	**4-Factor Model**
Intercept	0.007	-0.011	-0.013	-0.022
	(0.910)	*(0.858)*	*(0.854)*	*(0.756)*
Mispricing	-0.006	-0.006	-0.002	-0.001
	(0.263)	*(0.219)*	*(0.725)*	*(0.805)*
Size	0.003	0.004	0.005	0.004
	(0.589)	*(0.468)*	*(0.519)*	*(0.549)*
Idiosyncratic risk	-0.833*	-0.395	-0.820*	-0.315
	(0.066)	*(0.336)*	*(0.097)*	*(0.566)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.104	0.119	0.110	0.107
Observations	523	523	523	523

Appendix F.

*Table 15: Results of the regression of post-addition returns over the 30 days following the event window, on mispricing. Mispricing is estimated using P/E as of the last day of the event window. Size and idiosyncratic risk are taken from the day prior to the event window. P-values reported in parentheses, calculated using White's heteroskedasticity-consistent standard errors. * and ** indicate significance at the 10% and 5% levels respectively.*

	Adjusted Market Model	**Market Model**	**3-Factor Model**	**4-Factor Model**
Intercept	0.078	0.055	0.037	-0.019
	(0.456)	*(0.589)*	*(0.733)*	*(0.859)*
Mispricing	-0.025**	-0.021**	-0.018*	-0.018*
	(0.007)	*(0.016)*	*(0.055)*	*(0.061)*
Size	-0.002	-0.005	0.000	0.000
	(0.835)	*(0.627)*	*(1.000)*	*(0.963)*
Idiosyncratic risk	-2.838**	-2.077**	-3.059**	-1.681
	(0.006)	*(0.015)*	*(0.006)*	*(0.156)*
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.155	0.147	0.159	0.156
Observations	523	523	523	523