THERE IS SUCH A THING AS A FREE LUNCH, BUT IT WILL COST YOU

A STUDY ON COSTLY ARBITRAGE AND THE INDEX EFFECT

Sandra Andersson

Nicklas Kövamees

Bachelor Thesis

Stockholm School of Economics

2019

There is such a thing as a free lunch, but it will cost you: a study on costly arbitrage and the index effect

Abstract:

This thesis examines how stock returns around inclusions into and exclusions from the Swedish OMX Stockholm Benchmark and OMX Stockholm 30 indices are affected by arbitrage risk and transaction costs. We observe that during the 30 trading days preceding the announcement of a stock's inclusion in either index, stocks with high arbitrage risk and high transaction costs experience high positive abnormal returns, and vice versa. We conclude that arbitrage risk, i.e. the absence of close substitutes, and transaction costs both prevent arbitrage from flattening the demand curves for stocks. This suggests that mispricings are more likely to occur in stocks with high arbitrage risk and high transaction costs. However, we do not find any statistically significant corresponding relationships for stocks being excluded from either index. Finally, we study post-announcement abnormal returns and conclude that its relationship with arbitrage risk and transaction costs is vague in theory and ambiguous in practice.

Keywords:

Arbitrage risk, demand curve slopes, index effect, market efficiency, transaction costs

Authors:

Sandra Andersson (23838)
Nicklas Kövamees (23528)

Tutors:

Olga Obizhaeva, Assistant Professor, Department of Finance

Examiner:

[TBA]

Acknowledgements:

We thank our tutor Olga Obizhaeva, Assistant Professor at the Department of Finance, for highly valuable support and guidance. All errors are our own.

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics

Contents

APPENDICES

1. Introduction

There is much research confirming the tendency of a stock's price to increase when the stock is included in a major index, and to decrease when excluded. Multiple hypotheses have been presented for this phenomenon. Some predict permanent price effects: for example, that stocks are imperfect substitutes (cf. Wurgler & Zhuravskaya, 2002), that index constituents receive increased investor awareness (cf. Chen, Noronha and Singal, 2004), and that liquidity increases after a stock is included in an index (cf. Bechmann, 2004). One hypothesis also only predicts temporary price effects: that investors' – primarily index funds' – portfolio rebalancing causes temporary price pressure (cf. Harris and Gurel, 1986).

Some studies, e.g. Scholes (1972), Doukas et al. (2010), and Mendenhall (2004), suggest that arbitrage risk obstructs market efficiency, in the context of Fama's (1970) efficient market hypothesis. As such, arbitrage risk may entail that stock prices do not accurately reflect all available information. Wurgler and Zhuravskaya's (2002) studied the connection between arbitrage risk and the index effect. They looked at S&P 500 inclusions between 1976 and September 1989, concluding that stocks with high arbitrage risk experience large price gains upon inclusion in the S&P 500, and vice versa for stocks with low arbitrage risk. They explained this by concluding that high-arbitrage-risk are more sensitive to demand shocks because they have higher price elasticity.

We will look at how inter alia arbitrage risk affects stock prices at and around index inclusions and exclusions, together referred to as index events. Wurgler and Zhuravskaya (2002) correctly predicted that "the risk inherent in arbitrage between imperfect substitutes deters risk-averse arbitrageurs from flattening demand curves", and we will investigate whether this holds true also for two major indices on the Swedish stock exchange: the OMX Stockholm Benchmark index (OMXSB) and the OMX Stockholm 30 index (OMXS30). However, our sample will differ in a few notable aspects. In our sample the effective date – on which the pre-announced index changes take effect – follows 20 to 26 calendar days after the effective date. According to Beneish and Whaley (1996) the way stock prices react to index inclusions changed when pre-announcement was introduced for S&P 500 inclusions. They found that the total premium experienced upon inclusion increased: the immediate return upon announcement decreased from 3.7 to 3.1 percent with the new announcement policy, but however, they also found abnormal returns of an additional 4.1 percent between the announcement date and the effective date. This, they concluded, was an effect of index funds waiting until the effective date to start trading the stock. Another difference is that the S&P 500 and the Swedish indices has different methods for selecting constituents. Both the OMXSB and the OMXS30 have mechanical procedures for selecting index constituents and is possible to predict before the announcement date. The S&P 500 however has a selection committee and does not prevail what makes a stock to get picked for the indices. Therefore, different explanations might be of different importance for the respective indices.

In addition to idiosyncratic risk, some studies, including Pontiff (2005) and Doukas et al. (2010), have studied the effect of transaction costs (such as bid-ask spreads and price impact) and holding costs (such as idiosyncratic risk) on mispricings. They concluded that transaction costs and holding costs have a deterrent effect on arbitrageurs' ability to completely correct mispricings: both types of costs force arbitrageurs to take smaller and fewer positions than they otherwise would have, allowing mispricings to continue to exist.

Our study builds upon Wurgler and Zhuravskaya's (2002) research to (i) also examine the impact of transaction costs and (ii) examine whether their findings are applicable the Swedish stock market. As such, we formulate the following research question:

> To what extent do arbitrage risk, transaction costs, demand shock size, and heterogeneity of non-arbitrageurs' beliefs affect demand curve slopes on the Stockholm Stock Exchange?

2. Previous literature and theoretical framework

2.1. Theories on the index effect

There is plenty of previous research on the price effects that stocks experience when included or excluded from a major index. Most studies of the S&P 500, e.g. Shleifer (1986), Harris and Gurel (1986), and Chen, Noronha, and Singal (2004), show significant positive (negative) price effects when stocks are included in (excluded from) the index. Principally all studies find that some effect exists, but researchers disagree on why the effects occur and whether they are permanent or temporary. In the following sections, we will discuss the different theories about why index effects exist.

The Imperfect Substitutes Hypotheses

Wurgler & Zhuravskaya (2002) wrote that "In textbook theory, demand curves for stocks are kept flat by riskless arbitrage between perfect substitutes. However, individual stocks do not have perfect substitutes". For this reason, individual stocks can have positive impact on investors' diversification. The absence of perfect substitutes entails that stock prices will react to sudden demand shocks, positive or negative, such as the change in demand from index funds that follows index events. This forms the basis of the Imperfect Substitutes Hypothesis (ISH).

Kaul, Mehrotra and Morck (2000) studied an index reweighting of the Toronto Stock Exchange 300 index that took place in 1996, caused by a redefinition of the free-float – an event they argue cannot have revealed any private information about any of the companies concerned. The index weights of 31 stocks was increased, and said stocks experienced statistically significant excess returns of 2.3 % during the event week, with no price reversal being observed afterwards.

Shleifer (1986) found a positive price effect for stocks included in the S&P 500 and that the effect is permanent. He found this to support the ISH and concluded that long-term demand curves for stocks indeed slope downward.

The Price Pressure Hypothesis

When stocks are included (excluded) from indices, large volumes of stocks are normally traded when substantial demand from index funds and other index-following investors suddenly appears (disappears). According to the Price Pressure Hypothesis (PPH), as discussed by e.g. Harris and Gurel (1986), passive liquidity providers and other investors who accommodate such shifts in demand curves must be compensated for the risk they take in considerably larger volumes of stock than they otherwise would have. According to the PPH, investors buying (selling) large volumes of a stock will entail immediate price increases (decreases), following which providers of liquidity will be compensated as the price of said stock increases (decreases) back to its "natural" level.

Scholes (1972) studied large-block secondary sales of stocks, and found that large block trades do indeed have temporary effects on stock prices in line with the predictions of the PPH.

The Information Signaling Hypothesis

The Information Signaling Hypothesis proposes that if a stock is announced to be included in an index, this reflects good news about the company's prospects. Likewise, exclusions predict the contrary. While several studies, including Jain (1987), Denis et al. (2002), and Dhillon and Johnson (1991) finds support for the Information Signaling Hypothesis in studies of the S&P 500, the hypothesis is arguably of very little relevance for our study because both the OMXS30 and the OMXSB constituent lists are determined by mechanical calculations using easily observable variables. Therefore, the performance of such calculations ought to reveal no new information to the market – in comparison, the S&P 500 uses a selection board without any hard inclusion criteria.

The Liquidity Hypothesis

The Liquidity Hypothesis predicts that index inclusions lower transaction costs due to the included stock becoming traded more frequently. Bechmann (2004) studied these effects on the Danish KFX index and found that trading volume increased following index inclusions, which in turn lead to lower bid-ask spreads and other transaction costs.

Chen, Noronha and Singal (2004) also predicted that being added to an index makes it easier for a company to raise additional capital, which leads to positive price effects.

The Investor Awareness Hypothesis

A study by Chen, Noronha and Singal (2004) on both inclusions and exclusions of S&P 500 found evidence of asymmetric results: a prominent price increase takes place upon inclusion, but no significant decline is observable for exclusions. They explain this by the investor awareness hypothesis: investors' awareness of a company in the indices increases upon addition but does not go away in the event of exclusion. Further, Denis et al. (2003) concluded that when a stock included in an index, the monitoring of the becomes more intensified and in turn makes the firm more efficient.

2.2. Arbitrage costs as inhibitor of market efficiency

The Efficient Market Hypothesis holds that any mispricing of stocks will immediately disappear due to riskless arbitrage. However, Pontiff (2005) concluded that stock market mispricings can continue to exist due to holding costs, including idiosyncratic risk, and transaction costs. These costs force arbitrageurs to only take limited positions in the stock, limiting the ability of arbitrage to correct mispricings. The impact of transaction costs and idiosyncratic risk on arbitrage is not a particularly very well-researched area. While there are theoretical reviews acknowledging that idiosyncratic risk is costly to arbitrage and hinders efficient markets, such as Pontiff (2005), Hirshleifer (2001), and Shleifer and Vishny (2012), there is somewhat limited empirical evidence on the matter.

In two studies, however, Doukas et al. (2010) and Mendenhall (2004) found idiosyncratic risk to be a major deterrent of arbitrage activity. They both concluded that because arbitrageurs are to at least some extend averse to idiosyncratic risk, mispricings can continue to exist in the market due to arbitrageurs' limited willingness to enter into large arbitrage trades and thereby correct the mispricings.

2.3. Arbitrage risk and the index effect

Wurgler and Zhuravskaya (2002) examined the effect that arbitrage, principally measured as the lack of close substitutes, has on demand curves. They drew three main conclusions: (i) no perfect or even close substitutes exist between stocks, (ii) price increases following index inclusions are higher when arbitrage risk is higher, and (iii) the size of price effect is correlated to the size of the index fund demands. This means that arbitrage risk prevents arbitrageurs from flattening demand curves in practice, inhibiting market efficiency.

They developed a theoretical model with two types of investors: non-arbitrageurs, with heterogeneous beliefs, and arbitrageurs, with a zero-net-investment constraint as well as homogeneous and correct beliefs about fundamental value. They assume arbitrageurs are averse to the idiosyncratic risk that arises in arbitrage trades, because only a limited set of arbitrage trades are available at any time. When aggregating the demand curves from the two types of investors they found four factors that theoretically should steepen demand curve slopes: (a) the level of the arbitrage risk, i.e. the absence of close substitutes, (b) the arbitrageurs' risk aversion (c) the degree of heterogeneity of belief amongst non-arbitrageurs, and (d) the number of arbitrageurs.

Wurgler and Zhuravskaya (2002) estimated the demand shock size and arbitrage risk for each stock added to the S&P 500 index between 1976 and 1989. They estimated arbitrage risk as the historical variance of a portfolio consisting of (i) a long position in the included stock, (ii) a variable short position in substitutes to said stock, and (iii) a variable position in the risk-free rate. Using this model, they were able to conclude that high-arbitrage-risk stocks experience higher price increases upon inclusion to the S&P 500.

2.4. Index methodologies

The OMX Stockholm Benchmark index, OMXSB, includes around 90 of the companies with highest turnover on the Stockholm Stock Exchange. The OMX Stockholm 30 index, OMXS30, is comprised of the 30 most traded companies listed on the Stockholm stock Exchange. The constituent lists for both indices are revised twice a year, and revisions are announced – in our sample – between 20 and 26 calendar days before they are effective. Approximately, this translates to between c. 15 and 19 trading days.

The selection process for the OMXSB and OMXS30 is based on a set of mechanic criteria, detailed below, which arguably entails that no new information is revealed to the market upon an inclusion or exclusion announcement. As such, the information signaling hypothesis, cf. Bechmann (2004) can be ruled out as possible explanation for effects on the stocks in the Swedish Indices.

Nasdaq (2018) describes the constituent selection process for the OMXSB as follows: all shares are categorized into ICB supersectors, according to the industry the companies are active in. Within each supersector, shares are selected in order of decreasing free-float[1] market capitalization until 85% of the supersector's total market capitalization is reached. Following this, in addition any share in the top 10% of last twelve months' turnover amongst all listed shares is also included, but at least 25 shares. Additionally, shares in the bottom 30% of last twelve months' turnover amongst all listed shares are removed. The shares remaining at the end of this process will constitute the OMXSB.

For the OMXS30, Nasdaq (2016) specifies the following: the selection process is based on the SEK volume traded in each share during a six-month period. For a new stock to take the place of an old stock the new stock's traded volume must be amongst the 15 highest, and for an old company to get excluded and replaced by a new company, the old stock needs to fall out of the 45 highest-volume stocks.

This difference in methodology between the two indices leads to the OMXSB replacing a greater proportion of its constituents stocks every semi-annual review than the OMXS30. This is because for the OMXSB, the criteria for a non-constituent to become included and for a current constituent to stay included are the same. In the OMXS30, however, what is essentially a safety net makes it easier for stocks to remain included than for new stocks to become included. Therefore, fewer constituents are changed at each review.

[1] Free float is the proportion of the shares of a security which are freely available for trading on the market. Free-float excludes government holdings, controlling shareholders, company insider stakes and crossholdings. (Nasdaq, 2018)

3. Hypotheses

According to the model developed by Wurgler and Zhuravskaya (2002), arbitrage risk, demand shock size, and analyst target price dispersion are principally inhibitors of market efficiency. The same applies for bid-ask spreads, which we use as proxy for transaction costs. As such, all pre-announcement abnormal returns should increase in magnitude as our four dependent variables increases: this should apply for inclusions and exclusions, for positive and negative price effects, whether they are permanent or temporary, and for all event windows.

For post-announcement returns, however, the theoretical relationship between abnormal returns and our dependent variables is ill-defined. While price pressure effects should be amplified by our four dependent variables, the corresponding relationships for the other hypotheses that predict permanent price effects are vague. However, bearing in mind i.a. Doukas et al.'s (2010) and Mendenhall's (2004) findings on arbitrage costs and market inefficiencies, we hypothesize that all our dependent variables might also cause post-announcement abnormal returns to increase in magnitude.

Because much of previous research has focused primarily on index inclusions, we separate our hypotheses for index inclusions and exclusions. Therefore, we formulate the following hypotheses:

> **H_1**: for index inclusions, any abnormal returns leading up to and including the announcement day will increase in magnitude as arbitrage risk, demand shock size, analyst dispersion, and transaction costs increase.
>
> **H_2**: for index inclusions, any post-announcement abnormal returns, permanent or reverting will increase in magnitude as arbitrage risk, demand shock size, analyst dispersion, and transaction costs increase.
>
> **H_3**: for index exclusions, any abnormal returns leading up to and including the announcement day will increase in magnitude as arbitrage risk, demand shock size, analyst dispersion, and transaction costs increase.
>
> **H_4**: for index exclusions, any post-announcement abnormal returns, permanent or reverting will increase in magnitude as arbitrage risk, demand shock size, analyst dispersion, and transaction costs increase.

We will test our hypotheses for both the OMXSB and the OMXS30 indices.

4. Methodology

4.1. Model specification

Our models are largely set up like that of Wurgler and Zhuravskaya (2002), albeit with a few changes. They are all structured in a very similar way: they try to explain Cumulative Abnormal Returns in a set of different event windows, using different combinations of dependent variables.

The dependent variables used are summarized in **Table 1**. More details on their calculation can be found in the following sections. All estimation windows are specified as calendar time relative to each index event's Announcement Date (AD).

Table 1. Definitions of dependent variables and estimation windows

Variable	Definition	Estimation Window
A_1	Historical residual variance in a zero-net-investment arbitrage portfolio based on the event stock and a (typically short) position in the OMXS30	(-1825, -90) calendar days relative AD
A_2	Historical residual variance in a zero-net-investment arbitrage portfolio based the event stock and (typically short) positions in five substitute stocks	(-1825, -90) calendar days relative AD
A_3	Historical residual variance in a zero-net-investment arbitrage portfolio based the event stock and positions in the OMXS30 as well as five substitute stocks	(-1825, -90) calendar days relative AD
Demand shock size	Net assets in index funds following the relevant index, expressed as fraction of said index's total market capitalization	Last month end per 90 calendar days prior to AD
Analyst dispersion	Standard deviation of sell-side equity analysts' target prices for the event stock divided by the average	90 calendar days prior to AD
Amihud illiquidity	Average of daily absolute price change divided by daily SEK volume traded	(-210, -90) calendar days relative AD

4.1.1. Time measurement conventions

Our models contain a mixture of calendar day and trading day references, using the following convention: all our estimation windows use calendar day references, and all our event windows use trading day references. This is because we find that calendar days are, generally, have the best representation of the distance in time between two events. However, in the case of our event windows, it would be grossly impractical to use calendar days. Each stock is not traded two days every week, meaning that the sample in our study would rotate over time in seven-calendar-day cycles if we would use calendar days. To keep our sample constant over time (i.e. time relative to each Announcement Date) we use trading days for time measurement in our event windows.

4.2. Dependent variable: abnormal returns

To be able to draw reliable conclusions about the role of arbitrage at and around index events, we, like MacKinlay (1997) suggests, create a market-based model for stocks' expected return. The difference between the actual observed return and the expected return is the stock's Abnormal Return, *AR*, which will be the subject of our analysis.

We have modelled each stock's expected return as a version of a the four-factor model developed by Carhart (1997):

$$E(r_{i,t}) = r_{f,t} + \hat{\alpha}_i + \hat{\beta}_{RMRF,i} \times RMRF_t + \hat{\beta}_{SMB,i} \times SMB_t + \hat{\beta}_{HML,i} \times HML_t + \hat{\beta}_{MOM,i} \times MOM_t + \varepsilon_{i,t}$$

where *r* is logarithmic return and all factors are logarithmic.

For each stock we have conducted multiple linear regression with the stock price as independent variable and the Swedish House of Finance's Carhart factor data as dependent variables. In line with MacKinlay's (1997) recommendations, we want to avoid overlap between the estimation window and the event window. Also, in a compromise between sample size and beta instability over time, in line with Fraser and Groenewold (2000), Shanken (1992), and Roenfeldt, Griepentrog and Pflaum (1978)), we conduct this regression starting five years before each AD. As such, we use the window (-1825, -90) calendar days relative AD for our Carhart regression.

Following this, we calculate each stock's abnormal return on day *t* as the difference between actual observed return and the predicted return of our Carhart model:

$$\widehat{AR}_{i,t} = r_{i,t} - E(r_{i,t})$$

where *r* is logarithmic return.

4.3. Event windows

We define a set of event windows to analyze index events in various timeframes. As outlined in section 4.1.1, we use trading days as our measure of time.

For each event window, as specified in in **Table 2** below, we calculate Cumulative Abnormal Returns (CAR) as:

$$\widehat{CAR}_{i,t_1,t_2} = \sum_{t=t_1}^{t_2} \widehat{AR}_{i,t}$$

where t_1 and t_2 are the times of the Stockholm Stock Exchange's closing call on the first and last dates of the event window, respectively.

Table 2. Summary of event windows

Period (trading days relative AD)	Timing relative announcement
(-30, 0)	Leading up to and including announcement
(0, 15)	Following announcement
(15, 30)	Following announcement

Note: Announcement date (AD) is the date on which Nasdaq announced the stock's upcoming inclusion in or exclusion from the OMX Stockholm Benchmark or OMX Stockholm 30 index, whichever is applicable.

4.4. Arbitrage risk measures

We use Wurgler and Zhuravskaya's (2002) approach to estimating arbitrage risk: as the historical variance of a portfolio holding (i) a long (short) position in one unit of the stock which is included in or excluded from the OMXSB, referred to as the event stock, (ii) a short (long) position in a portfolio of substitutes, and (iii) a position in the risk-free rate. The position in the portfolio of substitutes is optimized so to minimize the overall portfolio's residual variance, and the position in the risk-free rate is set so to make the overall portfolio zero-net-investment. These measures of arbitrage risk attempt to capture the portion of risk that an arbitrageur is exposed to when speculating in a stock that cannot be hedged using short positions in close substitutes, i.e. idiosyncratic risk. This entails that our measures of arbitrage risk are equally applicable for arbitrageurs wanting to take long as well as short positions in the event stock.

Our first two arbitrage risk measures are very closely based on Wurgler and Zhuravskaya's (2002) measures of arbitrage risk. It is interesting to note that Wurgler and Zhuravskaya's (2002) method have been used for other topics than index effects as well, e.g. by Mendenhall (2004) in post-earnings announcement drift. These two first measures use two plausible substitutes to the event stock: (1) a short (long) position in the market portfolio, represented by the OMXS30 index, and (2) a portfolio of short (long) positions a small, select set of other stocks. Occasionally, (2) consists of long as well as short positions in a set of other stocks. We use the OMXS30 because it is principally the only Swedish stock index for which products such as futures are widely available and easily tradeable.

4.4.1. A_1 – Arbitrage risk based on the OMXS30 index

Calculating the first arbitrage risk measure, A_1, is relatively straightforward. For each index event, we want to estimate the event stock's return as a function of the OMXS30 index's return and the funding cost for this portfolio. If we did this using OLS regression over the previously-used estimation window (-1825, -90) calendar days relative AD – the same period as we use for estimating our Carhart factors – would result in the following equation:

$$\hat{r}_{i,t} = w_{i,OMXS30} \times r_{OMXS30,t} + w_{i,r_f} \times r_{f,t}$$

where r is logarithmic return, w denotes portfolio weights for OMXS30 and the risk-free rate, respectively, and $w_{i,OMXS30} + w_{i,r_f} - 1 = 0$ since the portfolio must be self-financing.

However, like Wurgler and Zhuravskaya (2002) noted, it is more convenient to use returns over the risk-free rate as opposed to raw returns because any set portfolio weights will yield a zero-net-investment portfolio; we are not required to explicitly constrain the sum of the portfolio weights to zero. As such, for each index event, we conduct OLS regression estimating the event stock's return over the risk-free rate as a function of the OMXS30 index's return over the risk-free rate. We do this again over our estimation window (-1825, -90) calendar days relative AD:

$$(\hat{r}_{i,t} - r_{f,t}) = w_{i,OMXS30} \times (r_{OMXS30,t} - r_{f,t})$$

where r is logarithmic return and $w_{i,OMXS30}$ is the portfolio weight in the OMXS30

The residual variance from this regression, A_1, is our first estimate of arbitrage risk.

4.4.2. A_2 – Arbitrage risk based on a portfolio of substitute stocks

The second arbitrage measure, A_2, is closely related to our first measure, A_1. However, instead of using the OMXS30 index as substitute, we instead use a portfolio of substitute stocks.

Wurgler and Zhuravskaya (2002) used the three closest stocks resulting from a process of ordering other stocks in the same industry by similarity in market capitalization and in market-to-book ratio, with the motivation that "solving for optimal zero-net-investment portfolio weights on each of thousands of assets for each of 259 stocks is too difficult both for us and for D. E. Shaw [a US investment management firm engaged in inter alia pair trading of stocks]". While doing exactly this ought to be largely unfeasible, modern computing has come a long way since 2002. As such, we select our substitute stocks by, for each event stock, computing the correlation against every stock in the Swedish House of Finance's FinBas database, which contains data on all stocks listed on the Stockholm Stock Exchange since 1979. We use daily returns data over the estimation window (-1825, -90) calendar days relative AD, and then sort the stocks by descending absolute correlation.

Any attempt to hedge a position in a stock by taking a short position in a different share class emitted by the same company would arguably be an ineffective hedge. This is because the two share classes are so close substitutes that any demand shock on one of the inevitably ought to spill other onto the other share class. For this reason, we remove any other share classes emitted by the same company from our sorted list. We also remove all shares that had been de-listed 90 calendar days before AD.

The transaction costs involved in short sales of many different stocks prohibit complex hedging portfolios. As such, we limit hedging portfolio complexity by selecting the top five shares according to the above sort by descending absolute correlation for use in our arbitrage portfolio. This entails a larger portfolio than the three-substitute-portfolio we used by Wurgler and Zhuravskaya (2002), a difference we believe is reasonable given that transaction costs have decreased considerably since the index events the studied that took place between 1976 and 1989, cf. Evans (2012).

Using these five select stocks, we perform the same type of regression as with the OMXS30 index above: for each event stock we conduct OLS regression over (-1825, -90) calendar days relative AD to estimate the index stock's return over the risk-free rate as a function of that of the other five substitute stocks:

$$(\hat{r}_{i,t} - r_f) = \sum_{j=j}^{n} w_{i,j} \times (r_{j,t} - r_{f,t})$$

where r is logarithmic return, j through n are the five selected substitute stocks, and $w_{i,j}$ denotes portfolio weights in substitute stocks j through n, respectively.

The residual variance, A_2, is our second estimate of arbitrage risk.

4.4.3. A_3 – Arbitrage risk based on the OMXS30 index in combination with a portfolio of substitute stocks

Arbitrage measures A_1 and A_2 are calculated closely following Wurgler and Zhuravskaya (2002)'s methodology. However, we believe that real arbitrageurs are likely to use a combination of a broad hedge and a narrower substitute stock hedge. Therefore, we have developed a third arbitrage risk estimate, A_3, which represents a combination of A_1 and A_2. It uses both the OMXS30 and the five substitute stocks selected using the methodology of A_2 as the contemplated substitutes.

For each event stock we conduct OLS regression over (-1825, -90) calendar days relative AD to estimate the stock's return over the risk-free rate as a function of that of the OMXS30 and the five substitute stocks:

$$(\hat{r}_{i,t} - r_f) = w_{i,OMXS30} \times (r_{OMXS30,t} - r_{f,t}) + \sum_{j=j}^{n} w_{i,j} \times (r_{j,t} - r_{f,t})$$

where r is logarithmic return, j through n are the five selected substitute stocks, and w_i denotes portfolio weights in the OMXS30 and substitute stocks j through n, respectively.

The residual variance from this regression, A_3, is our third estimate of arbitrage risk.

4.5. Demand shock size

We define *demand shock* as the portion of an included (excluded) firm's shares that investors will wish to buy (sell) only due to the stock being included in the relevant index. Intuitively, whenever an asset has positive price elasticity (as would under almost all conceivable circumstances be expected for stocks), a larger positive (negative) demand shock will entail a proportionally larger price increase (decrease).

This demand shock can principally be split up into two components: (i) demand from price inelastic index funds, and (ii) demand from other investors that either use the OMXSB as a benchmark or reference point. We believe (i) is very easy to estimate, whereas (ii) is very difficult to estimate: much like Wurgler and Zhuravskaya (2002) we only deem it practically possible to estimate the index fund component of the demand shock. For index events in the OMXSB, we estimate *demand shock* for an index event taking place at time *t* as:

$$demand\ shock_t = \frac{total\ net\ assets\ of\ index\ funds\ tracking\ the\ OMXSB_t}{total\ market\ capitalization\ of\ all\ shares\ in\ the\ OMXSB_t}$$

This is done using monthly data, and to avoid any overlap between our (one-day) estimation window and our event windows we use demand shock data per the last month-end as of 90 calendar days before AD.

For the OMXS30, we estimate *demand shock* analogously, using total net assets of index funds tracking the OMXS30 and the total market capitalization of all shares in the OMXS30.

It is important to distinguish *demand shock* from the proportion of shares held by index funds before, around, or after an index event. Many constituents of for example the OMXSB are also members of other indices that are closely followed by index funds, such as the OMXS30 and the SIXRX. Since the holdings of funds tracking other indices remain unaffected when a stock – all else equal – is included in or excluded from the OMXSB or the OMXS30, only the demand from index funds tracking the relevant index is of interest in this example.

4.5.1. Extrapolating missing values

We have only been able to retrieve total OMXSB market capitalization data starting December 2006, and for the OMXS30 starting February 2008.Therefore, we have extrapolated earlier market capitalization values as a function each respective index's closing values at the same time. For the OMXSB this becomes:

$$OMXSB\ \widehat{market\ capitalization}_t = A_{OMXSB} \times OMXSBPI\ closing\ price_t + B_{OMXSB}$$

where *A* and *B* are constants. For the OMXS30, our estimation is done analogously:

$$OMXS30\ \widehat{market\ capitalization}_t = A_{\mathrm{OMXSB}} \times OMXS30\ closing\ price_t + B_{\mathrm{OMXS30}}$$

Since the OMXSBPI is a price index tracking the value of all shares in the OMXSB, the OMXSBPI closing values and the market capitalization of the OMXSB are extremely correlated ($\rho = 0.997$). The same applies for the OMXS30, for which the corresponding correlation is also extremely high ($\rho = 0.999$). Therefore, we believe these extrapolations to be accurate.

4.6. Heterogeneity of non-arbitrageurs' beliefs as analyst dispersion

According to Wurgler and Zhuravskaya's (2002) model, heterogenous beliefs about a stock's value amongst non-arbitrageurs introduces a natural slope to demand curves. Their study estimates heterogeneity as the standard deviation of sell-side equity analysts' one-year-ahead earnings per share (EPS) forecasts, divided by the mean. However, their study did not find that analyst EPS dispersion had any statistically significant impact, instead concluding that "it seems likely that the weak effect of analyst [EPS] dispersion is due to its theoretically ambiguous relationship to true heterogeneity".

To find an estimate that is more closely related to true heterogeneity, we look at the dispersion of sell-side equity analysts' target prices rather than EPS forecasts. We believe this better represents true heterogeneity since it measures beliefs about the firm's fundamental value, rather than its performance in a single time period.

From Thomson Reuters we retrieve individual equity analysts' target prices for each event stock as per 90 calendar days before AD. Much like Wurgler and Zhuravskaya (2002), we disregard cases where target prices are only reported for fewer than five equity analysts. Using this data, we estimate heterogeneity for event stock *i* as:

$$Analyst\ dispersion_i = \frac{\sigma_{analyst\ target\ prices_i}}{\mu_{analyst\ target\ prices_i}}$$

4.7. Amihud illiquidity as proxy for transaction costs

In addition to the dependent variables in Wurgler and Zhuravskaya's (2002) study, we want to capture the transaction costs involved in arbitrage trades. Previous research has concluded that most measures of transaction costs in the stock market – i.e. primarily bid-ask spreads and market impact – are strongly related to each other, as studied by e.g. de Jong, Nijmanjong and Röell (1995), Amihud and Mendelson (1986), and Stoll (1989). Since these studies have also found that the largest component of transaction costs typically is market impact, we estimate market impact using the illiquidity measure developed by Amihud (2002). This measure is strongly related to one of the most well-known measures of market impact: Kyle's (1985) lambda.

Amihud's (2002) illiquidity measure is calculated by, for each day in a period, calculating the product of a stock's absolute return and the stock's trading volume in SEK for the same day. These products are then averaged over, in our case, (-210, -90) calendar days relative AD:

$$Amihud\ illiquidity = \frac{1}{D_i} \times \sum_{t=1}^{D_i} \left|r_{i,t}\right| \times SEK\ trading\ volume_{i,t}$$

Where D_i is the number of trading days for stock *i* in the relevant period and *r* is logarithmic return.

Because transaction costs can vary more over time than e.g. betas, cf. Choi, Salandro and Shastri (1988), we use a much shorter estimation window for Amihud illiquidity than for our other dependent variables.

5. Data

We look at the OMXSB, which as of May 2019 consists of 95 of the largest and most traded stocks on the Stockholm Stock exchange, selected according to a specific set of criteria that will be discussed later, and the OMXS30, which represents the 30 most traded stocks on the Stockholm Stock Exchange. While the OMXS30 has a fixed number of constituents, the OMXSB is an open-ended index whose number of constituents is not predetermined. For both the OMXS30 and the OMXSB index revisions are carried out biannually, and such index revisions are announced normally slightly less than a month in advance.

5.1. Sample selection and description

In arriving at a set of index events for us to analyze, we much like Wurgler and Zhuravskaya (2002) exclude index events related to M&A activity, bankruptcies, IPOs, and corporate demergers. This aims to control for index events that are caused by exogenous events, which may reveal information about the underlying stock and therefore distort our data.

Since index changes are preannounced, we need to know and the announcement dates – which are only available in the original press releases – in addition to the effective dates – which are easily accessible via e.g. Thomson Reuters). For this reason, we are constrained to index events announced in or after June 2003 for the OMXSB, and in or after December 1999 for the OMXS30, as we have been unable to retrieve earlier press releases. Because we need good Carhart factor data to be able to calculate abnormal returns we also cannot consider index events that have taken place in 2017 and later, since the Swedish House of Finance's Carhart factor data is only available up until the first few days of 2017.

Subject to the above two constraints, we have identified 143 inclusions and 130 exclusions for the OMXSB, between June 2003 and November 2016. For the OMXS30, we have identified 19 inclusions and 13 exclusions, between December 1999 and December 2015.

The large difference in number of observed index events between the two indices is caused by the OMXSB having considerably more constituents and by the OMXSB's constituent selection mechanism changing a larger fraction of its constituents at any single index revision, as outlined in the theoretical background.

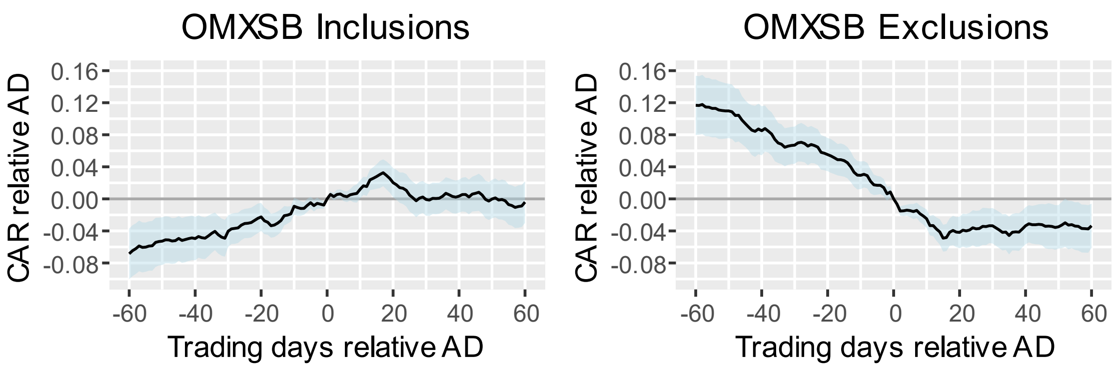
It is interesting to note that the number of index events we have found is somewhat higher than a previous study on the OMXSB: Karlsson and Koria (2018) looked at 87 inclusions and 79 exclusions between May 2009 and November 2016, which is roughly half the sample size that we have found. We believe that this difference exists because we have been able to retrieve older press releases dating back to 2003.

5.2. Summary statistics

5.2.1. Abnormal returns

In our sample, we find clear evidence of price increases around inclusions and price drops around exclusions. Looking at **Figure 1**, it is interesting to note that for OMXSB inclusions the price increase that proceeds AD does not seem to revert, whereas the price increase experienced the first 10-20 trading days after AD seems to revert at least partially within c. 30 trading days. Meanwhile, for exclusions, the price drops experienced both before and after AD do show any clear tendency of reversion.

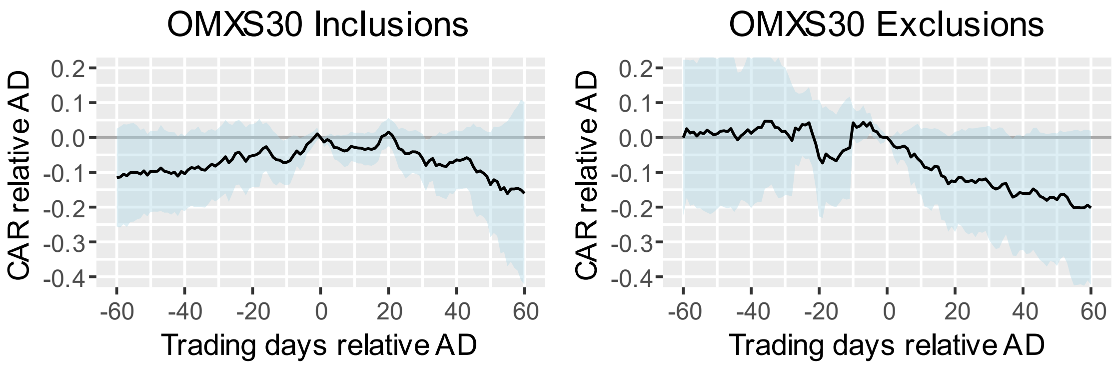
Figure 1. Cumulative Abnormal Returns around OMXSB index events



Note: Cumulative Abnormal returns around inclusions to and exclusions from the OMXSB index. Shaded blue areas represent 95% confidence intervals. Cumulative Abnormal Returns are calculated as actual returns over expected returns from a Carhart four-factor model.

Looking at the OMXS30, depicted in **Figure 2**, none of the apparent price patterns around index events are significant and it is hard to draw any reliable conclusions. At least partially, the small sample size for the OMXS30 ought to be to blame for this.

Figure 2. Cumulative Abnormal Returns around OMXS30 index events



Note: Cumulative Abnormal returns around inclusions to and exclusions from the OMXSB index. Shaded blue areas represent 95% confidence intervals. Cumulative Abnormal Returns are calculated as actual returns over expected returns from a Carhart four-factor model.

In **Table 3** follows an overview of CAR across our event windows. All event windows are specified as trading days relative the announcement date. We find statistically significant cumulative abnormal returns at the 0.1% level in all event windows for OMXSB inclusions, and in the (-30, 0) and (0, 15) event windows for OMXSB exclusions. In the (15, 30) event window the presence of abnormal returns is not significant at the 5% level. Moreover, we do not find any significant CAR in any event window for the OMXS30 – again p,ossibly due to the small OMXS30 sample size.

Table 3. Cumulative Abnormal Returns around index events by event window

Index	Type	Event window	N	Mean	P-value	St. Dev.	Min	Max
OMXSB	Inclusions	(-30, 0)	143	0.039	1.25e-04	0.12	-0.348	0.374
		(0, 15)	143	0.03	1.38e-04	0.092	-0.382	0.226
		(15, 30)	143	-0.029	2.44e-05	0.078	-0.294	0.215
	Exclusions	(-30, 0)	130	-0.069	2.91e-08	0.133	-0.504	0.366
		(0, 15)	130	-0.05	4.70e-08	0.12	-0.348	0.202
		(15, 30)	130	0.016	0.061	0.092	-0.382	0.331
OMXS30	Inclusions	(-30, 0)	19	-0.034	0.836	0.097	-0.275	0.098
		(0, 15)	19	0.075	0.116	0.212	-0.210	0.745
		(15, 30)	19	-0.026	0.280	0.109	-0.232	0.182
	Exclusions	(-30, 0)	13	-0.084	0.836	0.179	-0.485	0.161
		(0, 15)	13	-0.018	0.116	0.303	-0.411	0.779
		(15, 30)	13	-0.046	0.280	0.146	-0.457	0.156

Note: this table represents Cumulative Abnormal Returns for stocks being included in and excluded from the OMXSB and OMXS30 indices, in a set of trading day event windows around the inclusion or exclusion announcement. Cumulative Abnormal Returns are calculated as actual returns over expected returns from a Carhart four-factor model.

5.3. Arbitrage risk

Table 4 and **Table 5** below present summary statistics of arbitrage risk measures for OMXSB and OMXS30 index events, respectively. The explained variance is calculated as original variance minus the residual variance for each arbitrage measure, e.g. E1 is defined as Var(Ri, t) – A1. Because our arbitrage risk models do not include constants – they would have no real-world interpretation – we occasionally encounter negative values for explained variance. Additionally, we calculate fractions of explained variance over original variance.

Looking at explained variance divided by original variance, we can see that none of our arbitrage portfolios constitute a generally effective hedge for either index. The original variance, $Var(R_{i,t} - R_f)$, is on the same order magnitude as the residual variance of our hedged portfolios. However, it is likely that the arbitrage portfolios' residual variances contain a large fraction of idiosyncratic risk, which is priced lower by investors. At the same time the original variance, which contains a greater fraction of systemic risk, ought to be priced higher by investors, c.f. Brockman, Schutte and Yu (2009).

Table 4. Summary statistics for the arbitrage risk measures for OMXSB inclusions and exclusions

Variable	N	Mean	St. Dev.	Min	Max
A_1	273	0.000784	0.000779	0.000075	0.005436
A_2	273	0.000745	0.000757	0.000068	0.004952
A_3	273	0.000733	0.000742	0.000066	0.004930
$Var(R_{i,t} - R_f)$	273	0.000919	0.000847	0.000157	0.005989
E_1	273	0.000135	0.000119	-0.000002	0.000553
E_2	273	0.000174	0.000163	0.000001	0.001037
E_3	273	0.000186	0.000165	0.000006	0.001059
$E_1 / Var(R_{i,t} - R_f)$	273	0.167795	0.120279	-0.006373	0.677080
$E_2 / Var(R_{i,t} - R_f)$	273	0.218300	0.143241	0.001840	0.684526
$E_3 / Var(R_{i,t} - R_f)$	273	0.229369	0.139044	0.026430	0.710634

Note: A is the residual variance for each arbitrage portfolio, as specified in section 4.4. $Var(R_{i,t} - R_f)$ is the original variance. *E* is the explained variance for each arbitrage portfolio: original variance minus the residual variance. All variables have been estimated over (-1825, -90) calendar days relative each event's announcement date.

Table 5. Summary statistics for the arbitrage risk measures for OMXS30 inclusions and exclusions

Variable	N	Mean	St. Dev.	Min	Max
A_1	32	0.000909	0.001207	0.000131	0.004712
A_2	32	0.000733	0.000844	0.000098	0.003068
A_3	32	0.000719	0.000837	0.000097	0.003009
$Var(R_{i,t} - R_f)$	32	0.001103	0.001462	0.000202	0.006028
E_1	32	0.000194	0.000281	0.000006	0.001316
E_2	32	0.000370	0.000706	0.000005	0.003329
E_3	32	0.000384	0.000704	0.000011	0.003325
$E_1 / Var(R_{i,t} - R_f)$	32	0.198047	0.121480	0.018071	0.437274
$E_2 / Var(R_{i,t} - R_f)$	32	0.285715	0.153714	0.016240	0.615169
$E_3 / Var(R_{i,t} - R_f)$	32	0.305543	0.146183	0.035899	0.620368

Note: A is the residual variance for each arbitrage portfolio, as specified in section 4.4. $Var(R_{i,t} - R_f)$ is the original variance. *E* is the explained variance for each arbitrage portfolio: original variance minus the residual variance. All variables have been estimated over (-1825, -90) calendar days relative each event's announcement date.

Our three measures of arbitrage risk, A_1 through A_3, are highly correlated: for OMXSB, ρ_{A_1,A_2} = 0.9954, ρ_{A_1,A_3}= 0.9966, and ρ_{A_2,A_3}= 0.9996; for the OMXS30 ρ_{A_1,A_2} = 0.9671, ρ_{A_1,A_3}= 0.9712, and ρ_{A_2,A_3}= 0.9997. This is line with Wurgler and Zhuravskaya (2002), who found a correlation of c. 0.98 between their A_1 and A_2 measures. It is interesting to note that A_3, which hedges using the OMXS30 index in combination with five substitute stocks, only adds a limited reduction in residual variance in comparison with A_2, which hedges only using five substitute stocks.

5.3.1. Demand shock size, Amihud illiquidity and analyst dispersion

As illustrated in **Table 6** and **Table 7**, *Analyst dispersion* has considerably fewer observations than the other variables. This is due to data availability: for many index events, especially older, Thomson Reuters has no observations on analysts' target prices, or fewer than five observations which is our threshold. It is interesting to note that Amihud illiquidity is much lower for OMXS30 stocks than for OMXSB stocks. This is in line with our expectations and follows from the OMXS30's nature: the it is constructed to represent the 30 most liquid stocks on the Stockholm Stock Exchange.

Table 6. Summary statistics for dependent variables except arbitrage risk measures for OMXSB inclusions and exclusions

Variable	N	Mean	St. Dev.	Min	Max
Demand shock size	273	0.004955	0.003818	0	0.013650
Analyst dispersion	72	0.168669	0.124665	0.025764	0.757517
Amihud illiquidity	273	0.000028	0.000073	0.0000001	0.000679

Note: Demand shock size is net assets in index funds following the OMXSB as fraction of the total market capitalization of all stocks that constitute the OMXSB, per last month-end 90 calendar days before each event's announcement date. Analyst dispersion is the standard deviation of sell-side equity analysts' target prices divided by the mean, per 90 calendar days prior to the announcement date. Amihud illiquidity is the historical average of daily absolute price change divided by daily SEK volume traded over (-210, -90) calendar days relative the announcement date.

Table 7. Summary statistics for dependent variables except arbitrage risk measures for OMXS30 inclusions and exclusions

Variable	N	Mean	St. Dev.	Min	Max
Demand shock size	32	0.002115	0.002861	0	0.009858
Analyst dispersion	11	0.212840	0.113011	0.094915	0.419355
Amihud illiquidity	32	0.000001	0.000001	0.00000005	0.000003

Note: Demand shock size is net assets in index funds following the OMXS30 as fraction of the total market capitalization of all stocks that constitute the OMXS30, per last month-end 90 calendar days before each event's announcement date. Analyst dispersion is the standard deviation of sell-side equity analysts' target prices divided by the mean, per 90 calendar days prior to the announcement date. Amihud illiquidity is the historical average of daily absolute price change divided by daily SEK volume traded over (-210, -90) calendar days relative the announcement date.

It should also be noted that demand shock size is heavily dependent on time: **Table 8** shows how net assets in index funds following the OMXSB have increased considerably since 2003, as index funds have grown increasingly popular. The index was incepted only in July 2002, and up until August 2005 there were no index funds following the OMXSB. **Table 9** shows the corresponding trend for the OMXS30. Fund net asset data has been retrieved from Morningstar Direct, and OMXSB and OMXS30 market capitalization data has been retrieved from Thomson Reuters.

Table 8. OMXSB demand shock size over time.

	2003-12-31	**2008-12-31**	**2013-12-31**	**2018-12-31**
OMXSB market capitalization (SEK bn)	1,561.6	1,640.2	3,409.0	3,870.2
Net assets in OMXSB index funds (SEK bn)	0.0	4.6	23.5	69.1
Demand shock size	0.0000	0.0028	0.0069	0.0179

Note: OMXSB market capitalization is the market capitalization of all stocks that, at each point in time, constitute the OMXSB. Demand shock size is calculated as net assets in OMXSB index funds divided by OMXSB market capitalization.

Table 9. OMXS30 demand shock size over time.

	2003-12-31	**2008-12-31**	**2013-12-31**	**2018-12-31**
OMXS30 market capitalization (SEK bn)	1,556.9	1,589.0	3,347.3	3,535.9
Net assets in OMXS30 index funds (SEK bn)	0.0	6.0	22.1	26.7
Demand shock size	0.0000	0.0038	0.0066	0.0076

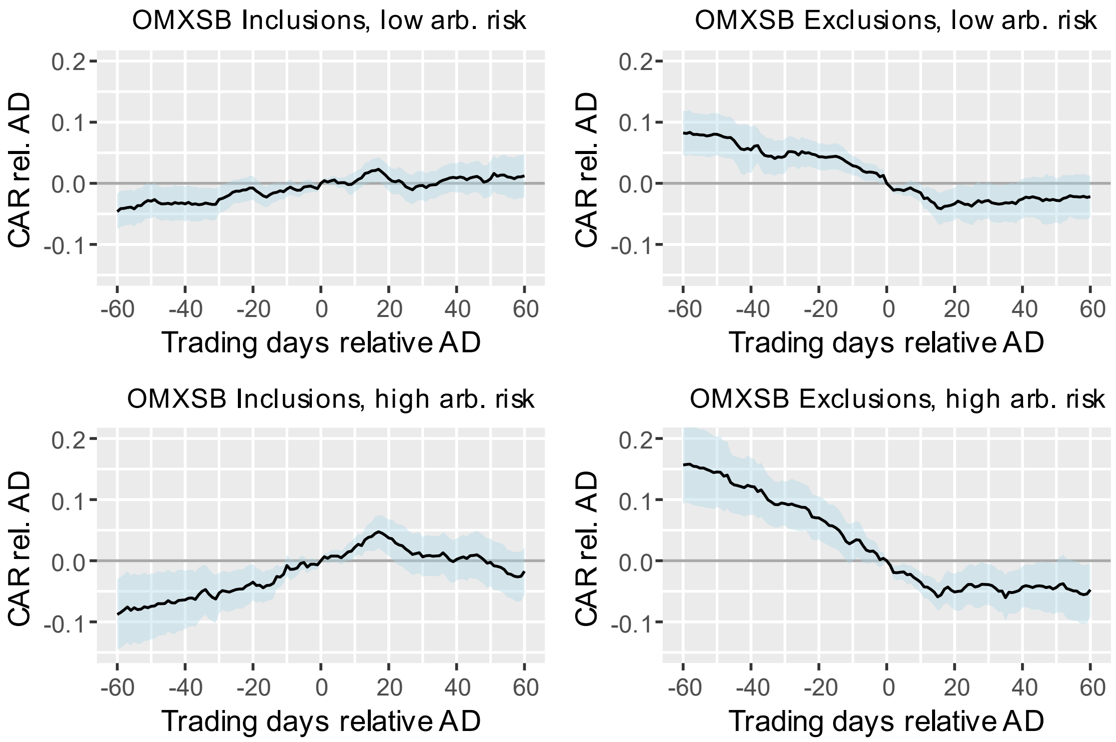
Note: OMXS30 market capitalization is the market capitalization of all stocks that, at each point in time, constitute the OMXS30. Demand shock size is calculated as net assets in OMXS30 index funds divided by OMXS30 market capitalization.

6. Empirical results

6.1. Graphical summary of results

Looking at **Figure 3**, we can see that the level of arbitrage risk, measured as A_1, affects pre-announcement abnormal returns for both inclusions and exclusions for the OMXSB: stocks with high arbitrage risk experience larger price gains (drops) before inclusion (exclusion) announcements. It is however not possible to visually identify any clear relationship between level of arbitrage risk and post-announcement abnormal returns.

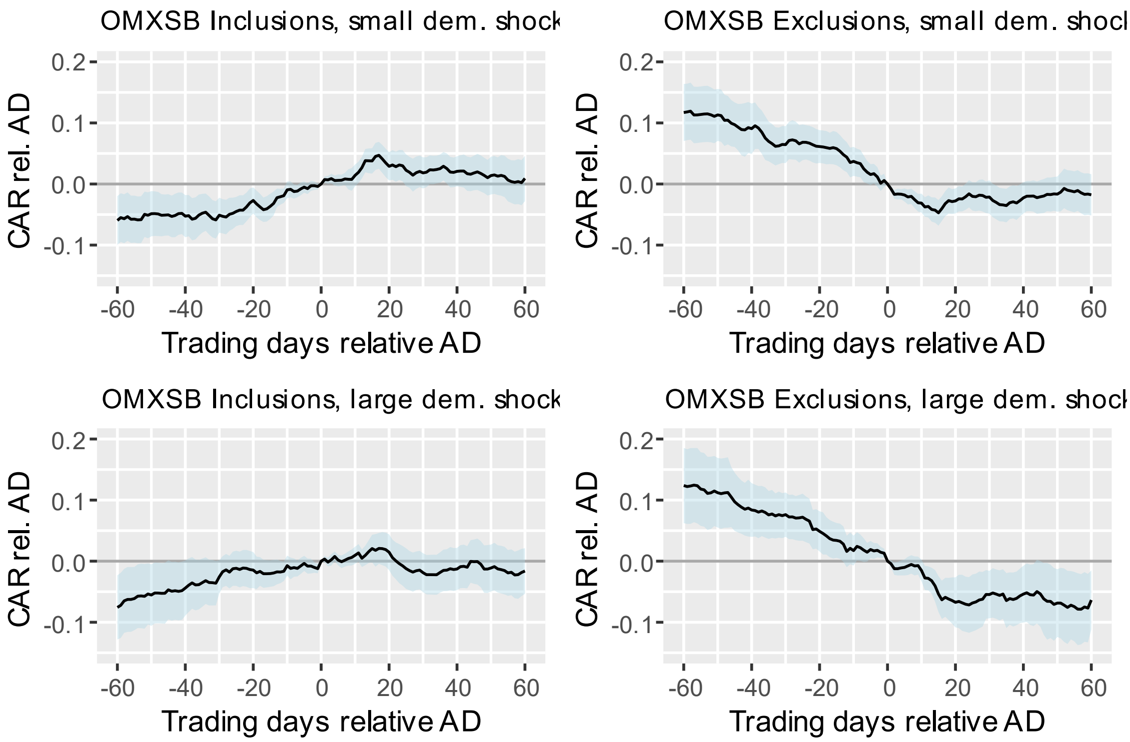
Figure 3. Cumulative Abnormal Returns around OMXSB index events, by arbitrage risk (A_1).



Note: These graphs represent Cumulative Abnormal Returns for stocks included in and excluded from the OMXSB, by level of arbitrage risk measured as A_1. Cumulative Abnormal Returns are calculated as actual returns over expected returns from a Carhart four-factor model. Shaded blue areas represent 95% confidence intervals. To control for variance in demand shock size, inclusions and exclusions with demand shock size in the top or bottom quartiles are removed. The remaining observations are split up into above-median and below-median arbitrage risk.

Figure 4 illustrates that for demand shock size, it is more difficult to visually discern any patterns. However, for exclusions, large demand shocks seem to entail permanent price drops whereas small demand shocks seem to entail only temporary effects.

Figure 4. Cumulative Abnormal Returns around OMXSB index events, by demand shock size.



Note: These graphs represent Cumulative Abnormal Returns for stocks included in and excluded from the OMXSB, by level of demand shock size. Cumulative Abnormal Returns are calculated as actual returns over expected returns from a Carhart four-factor model. Shaded blue areas represent 95% confidence intervals. To control for arbitrage risk, index events with arbitrage risk, measured as A_1, in the top or bottom quartiles are removed. The remaining observations are split up into above-median and below-median arbitrage risk.

Due to our small sample size for the OMXS30, it is not meaningful to produce any graphs corresponding to the above two. Instead, the OMXS30 will be investigated in the sections to follow.

6.2. Pre-announcement abnormal returns – OMXSB

Table 10 investigates the relationships between pre-announcement abnormal returns and our dependent variables in different combinations, for OMXSB inclusions. Specifications 1, 2 and 3 show that all three measures of arbitrage risk, A_1 through A_3, are indeed significant at the 5% level for pre-announcement abnormal returns for OMXSB inclusions. Specifications 4 and 5 show that demand shock size exhibits no significant linear relationship with pre-announcement abnormal returns. Specification 6 shows that explained variance does not have any effect on pre-announcement abnormal returns, in line with our expectations. Specification 7 illustrates that analyst dispersion also does not exhibit any statistically significant relationship with pre-announcement abnormal returns. Specifications 8 and 9 evaluate the effects separately of arbitrage-risk-demand-shock interaction and Amihud illiquidity. Specification 10 confirms that these two in combination plays a large role in explaining pre-announcement abnormal returns, with the arbitrage-risk-demand-shock interaction effect significant at the 5% level, and Amihud illiquidity at the 1% level.

Table 10. Regression of pre-announcement Cumulative Abnormal Returns for OMXSB Inclusions

	Dependent variable:									
	CAR (-30, 0) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	24.376**			26.597**		24.512*	-71.551		13.407	
	(11.964)			(12.805)		(14.317)	(70.353)		(12.472)	
A_2		25.032**			27.096**					
		(12.488)			(13.309)					
A_3			25.294**							
			(12.669)							
Analyst dispersion							0.045			
							(0.234)			
Demand shock size				1.249	1.149		-0.449			
				(2.520)	(2.510)		(4.996)			
E_1						-1.667				
						(95.601)				
A_1 * demand shock size								4,330.162		5,586.253**
								(2,660.711)		(2,589.170)
Amihud illiquidity									327.792**	410.154***
									(126.696)	(118.882)
Constant	0.020	0.021	0.021	0.012	0.013	0.020	0.053	0.027**	0.020	0.012
	(0.014)	(0.014)	(0.014)	(0.022)	(0.021)	(0.015)	(0.057)	(0.013)	(0.013)	(0.013)
Observations	143	143	143	143	143	143	37	143	143	143
Adjusted R^2	0.022	0.021	0.021	0.016	0.015	0.015	-0.051	0.011	0.060	0.082

Note: *p<0.1; **p<0.05; ***p<0.01

Standard errors in parentheses.

Note: This table illustrates OLS regression results for models of cumulative abnormal returns during the – (-30, 0) trading day period relative the announcement of stocks' addition to the OMXSB index. *A* represents our estimates of arbitrage risk. Analyst dispersion represents the standard deviation of sell-side equity analysts' target prices divided by the mean. E_1 is the explained variance of the A_1 arbitrage portfolio: original variance minus the residual variance. Demand shock size represents net assets in index funds following the OMXSB as fraction of the total market capitalization of all stocks that constitute the OMXSB. Amihud illiquidity represents the historical average of daily absolute price change divided by daily SEK volume traded.

Appendix A shows the corresponding regressions for OMXSB exclusions. It is interesting to note that the relationship between pre-announcement abnormal returns and essentially all evaluated variables seems weaker than for OMXSB Inclusions. The only statistically significant results are from specifications 8 and 10, which show that the interaction between arbitrage risk and demand shock size has a clearly significant impact on pre-announcement abnormal returns. While Amihud illiquidity had strong significance for OMXSB Inclusions, we do no find significant support for it in the case of OMXSB exclusions. Additionally, it should be noted that all specifications but one for OMXSB exclusions exhibit statistically significant constant, meaning that our model is missing some factor in this case.

6.3. Post-announcement abnormal returns – OMXSB

For OMXSB inclusions, our regressions in the post-announcement windows (0, 15) and (15, 30) trading days relative AD, respectively, can be found in **Appendix B**. Our models show very limited significance in this case: demand shock size seems significant in the (0, 15) event window, but only in combination with A_1 and A_2, which in the relevant specifications receive non-significant constants with negative signs – contrary to what is predicted by theory.

The corresponding regressions for OMXSB exclusions can be found in **Appendix C**. Our findings for exclusions principally correspond to those for inclusions: our models show very limited significance. The only interesting finding is that for the (15, 30) event window, specification 6 with A_1 in combination with E_1 produces significant results.

6.4. Pre-announcement abnormal returns – OMXS30

Table 11 below shows the above-used model specifications applied on OMXS30 inclusions in the (-30, 0) trading days relative AD window. Much like for OMXSB inclusions in the same event window, we see in specifications 1-3 that all three measures of arbitrage risk are significant: A_1 at the 5% level, and A_2 as well as A_3 at the 1% level. No other variable than arbitrage risk exhibits statistically significant results.

Table 11. Regression of pre-announcement Cumulative Abnormal Returns for OMXS30 inclusions

	Dependent variable:									
	CAR (-30, 0) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	110.697**			109.920**		102.082	-163.890		102.644**	
	(42.149)			(43.266)		(74.741)	(107.567)		(45.036)	
A_2		174.013***			173.664***					
		(52.160)			(54.473)					
A_3			175.475***							
			(52.659)							
Analyst dispersion							0.475			
							(0.260)			
Demand shock size				6.310	0.580		-18.685			
				(15.526)	(14.559)		(11.527)			
E_1						66.839				
						(472.018)				
A_1 * demand shock size								10,450.150		12,998.110
								(8,526.737)		(8,373.577)
Amihud illiquidity									51,011.420	136,772.800
									(85,407.570)	(89,212.150)
Constant	-0.021	-0.052	-0.051	-0.033	-0.053	-0.022	0.061	0.054	-0.040	-0.021
	(0.056)	(0.054)	(0.054)	(0.065)	(0.061)	(0.059)	(0.069)	(0.051)	(0.065)	(0.069)
Observations	19	19	19	19	19	19	7	19	19	19
Adjusted R^2	0.247	0.360	0.360	0.208	0.320	0.201	0.239	0.027	0.217	0.099
Note:									*$p<0.1$; **$p<0.05$; ***$p<0.01$*	
									Standard errors in parentheses.	

Note: This table illustrates OLS regression results for models of cumulative abnormal returns during the – (-30, 0) trading day period relative the announcement of stocks' addition to the OMXS30 index. *A* represents our estimates of arbitrage risk. Analyst dispersion represents the standard deviation of sell-side equity analysts' target prices divided by the mean. E_1 is the explained variance of the A_1 arbitrage portfolio: original variance minus the residual variance. Demand shock size represents net assets in index funds following the OMXSB as fraction of the total market capitalization of all stocks that constitute the OMXS30. Amihud illiquidity represents the historical average of daily absolute price change divided by daily SEK volume traded.

Appendix D shows the corresponding regressions for OMXS30 exclusions. Unlike for the two most comparable regressions, OMXS30 inclusions as well as OMXSB exclusions, we see very few significant results and in no specifications do all independent variables have significant coefficients.

6.5. Post-announcement abnormal returns – OMXS30

Interestingly, our models yield statistically significant for OMXS30 inclusions in the (0, 15) trading days relative AD event window, unlike our findings for OMXSB inclusions in the same event window. Specifications 1-3 of **Table 12** show that arbitrage risk does have a statistically significant impact on post-announcement abnormal returns in the (0, 15) event window form OMXS30 inclusions: A_1 at the 5%

level and A_2 and A_3 at the 1% level. The same is also true for arbitrage risk in combination with demand shock size, which is significant at the 1% level as illustrated in specification 8.

Table 12. Regression of post-announcement Cumulative Abnormal Returns for OMXS30 inclusions in (0, 15) trading days relative AD

	Dependent variable:									
	CAR (0, 15) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-51.147**			-49.694**		-103.044***	-256.356		-62.163***	
	(19.302)			(18.138)		(30.313)	(138.120)		(18.713)	
A_2		-80.529***			-74.884***					
		(23.823)			(23.285)					
A_3			-80.155***							
			(24.257)							
Analyst dispersion							-0.046			
							(0.333)			
Demand shock size				-11.805*	-9.370		-0.382			
				(6.509)	(6.224)		(14.802)			
E_1						402.658*				
						(191.435)				
A_1 * demand shock size								-10,850.150***		-10,626.300***
								(3,123.316)		(3,272.665)
Amihud illiquidity									69,778.510*	12,015.680
									(35,488.960)	(34,867.000)
Constant	0.010	0.024	0.023	0.034	0.040	-0.001	0.096	-0.013	-0.017	-0.020
	(0.025)	(0.025)	(0.025)	(0.027)	(0.026)	(0.024)	(0.088)	(0.019)	(0.027)	(0.027)
Observations	19	19	19	19	19	19	7	19	19	19
Adjusted R^2	0.251	0.367	0.355	0.340	0.411	0.376	0.372	0.381	0.359	0.347
Note:										**p<0.1; **p<0.05; ***p<0.01*
										Standard errors in parentheses.

Note: This table illustrates OLS regression results for models of cumulative abnormal returns during the – (0, 15) trading day period relative the announcement of stocks' addition to the OMXS30 index. *A* represents our estimates of arbitrage risk. Analyst dispersion represents the standard deviation of sell-side equity analysts' target prices divided by the mean. E_1 is the explained variance of the A_1 arbitrage portfolio: original variance minus the residual variance. Demand shock size represents net assets in index funds following the OMXSB as fraction of the total market capitalization of all stocks that constitute the OMXS30. Amihud illiquidity represents the historical average of daily absolute price change divided by daily SEK volume traded.

Our findings in the previous paragraph are similar for the (15, 30) trading days relative AD event window, as illustrated in **Appendix E**. Again, A_1 is significant at the 5% level and A_2 and A_3 at the 1% level. Unlike pre-announcement abnormal returns for OMXS30 inclusions, we also see in specification 8 that the arbitrage-risk-demand-shock interaction effect is significant at the 1% level.

Appendix F contains corresponding regression tables for OMXS30 exclusion post-announcement returns. In the (0, 15) trading days relative AD event window our findings are principally the same for OMXS30 exclusions as for OMXS30 inclusions, but with A_1 through A_3 significant at the 1% level. However, for OMXS30 exclusions in this event window we do not show any significance for demand shock size.

For the (15, 30) trading days relative AD event window, our models show limited significance, except for A_1 in specification 1 that is significant at the 5% level. Notably, neither A_2 or A_3 shows significant results even though A_1 does.

7. Discussion of Results

7.1. Theory and our findings on the index effect

While the presence or absence of index effects is not in itself the primary subject of our study, it is interesting to see how our findings on abnormal returns around index events differ from those of previous studies.

For OMXSB inclusions we have found that the positive pre-announcement abnormal returns are permanent and do not revert, whereas the post-announcement abnormal returns at least partially revert within 30 days. However, our data does not allow us to conclude that the full price increase experienced in the (0, 15) trading days window reverts within 30 trading days after AD. This means that, in the case of OMXSB inclusions, our study shows support for the imperfect substitutes hypothesis, the liquidity hypothesis, and the investor awareness hypothesis, as well as the price pressure hypothesis.

For OMXSB exclusions, we observe negative returns both pre-announcement and post-announcement, and our data does not support any reversion effects although it does not allow us to rule reversion effects out either. As such, for OMXSB exclusions our data shows support for the imperfect substitutes hypothesis, the liquidity hypothesis, and the investor awareness hypothesis, but does not show support for the price pressure hypothesis.

For the OMXS30, it is not possible for us to draw any corresponding conclusions due to the small sample sizes: our data neither supports nor disproves any of the theories.

7.2. Arbitrage risk, transaction costs, and the index effect

Our OMXSB models have consistently low R^2-values. However, we do not find this problematic since low R^2-values is a common occurrence in panel studies of stock returns over short time periods, cf. Wurgler and Zhuravskaya (2002). We also observe that our models for the OMXS30 have considerably higher R^2-values than our models of the OMXSB, which is somewhat puzzling. This phenomenon has no clear explanation in theory, although one possible explanation could be that this is at least partially caused by overfitting due to the very small sample size, cf. Hurvich and Tsai (1989).

We can also see that analyst dispersion does not show any significance in any context. Although our study looks at analysts' target prices, this is in line with the findings of Wurgler and Zhuravskaya (2002) who used at analysts' earnings per share forecasts and concluded that "It seems likely that the weak effect of analyst dispersion is due to its theoretically ambiguous relationship to true heterogeneity".

Additionally, we can note that Amihud illiquidity is clearly less significant for OMXS30 than OMXSB. We believe that this is not due to any fundamental difference in how illiquidity affects index events in the two indices, but rather because of the much-lower illiquidity in the OMXS30 than in the OMXSB, cf. **Table 6** and **Note***:* Demand shock size is net assets in index funds following the OMXSB as fraction of the total market capitalization of all stocks that constitute the OMXSB, per last month-end 90 calendar days before each event's announcement date. Analyst dispersion is the standard deviation of sell-side equity analysts' target prices divided by the mean, per 90 calendar days prior to the announcement date. Amihud illiquidity is the historical average of daily absolute price change divided by daily SEK volume traded over (-210, -90) calendar days relative the announcement date.

Table 7 on page 25. As discussed on the aforementioned page, this difference follows from the nature of the OMXS30: because it represents the 30 most traded stocks on the Stockholm Stock Exchange, its illiquidity is naturally very low and the very small effects of illiquidity difficult to observe.

7.2.1. OMXSB pre-announcement abnormal returns

Our results show that arbitrage risk has statistically significant impact on pre-announcement abnormal returns for OMXSB inclusions. However, arbitrage risk alone has no statistically significant impact on OMXSB exclusions. The interaction effect between arbitrage risk and demand shock size is statistically significant in combination with Amihud illiquidity for OMXSB inclusions and only by itself for OMXSB inclusions.

While our models yield results very much in line with Wurgler and Zhuravskaya's (2002) for our OMXSB inclusions, the fact that our models for OMXSB exclusions show statistically significant constants in most specifications means that our models ought to be missing some key characteristic that differentiates inclusions from exclusions in the (-30, 0) pre-announcement window.

Looking back at specification 10 of

Table 10, which is the specification with largest explanatory power for OMXSB inclusion pre-announcement abnormal returns measured as adjusted R^2, we find clear and statistically significant effects on pre-announcement abnormal returns for the interaction effect between arbitrage risk and demand shock size (significant at the 5% level) and Amihud illiquidity (significant at the 1% level). As such, in the case of OMXSB inclusion pre-announcement abnormal returns, we find support for our theoretical prediction that arbitrage risk and illiquidity both inhibit market efficiency.

7.2.2. OMXSB post-announcement abnormal returns

For OMXSB post-announcement abnormal returns our models show very limited significance and feature significant constants. This is not unexpected, given that theory does not provide any clear relationship between either dependent variable or post-announcement abnormal returns. While some specifications do yield significant coefficients for individual variables, the results are inconsistent and do not allow us to draw any conclusions about the effects of our dependent variables on OMXSB post-announcement abnormal returns.

What however is unexpected is that explained variance seems to have a strong negative relationship to abnormal returns in (15, 30) trading days relative AD for OMXSB exclusions, significant at the 1% level. This would entail that investors, in stark contrast to theoretical predictions and for some unknown reason, care about the degree of hedgeable risk in addition to hedgeable risk.

7.2.3. OMXS30 pre-announcement abnormal returns

Arbitrage risk alone shows statistically significant results for OMXS30 inclusion pre-announcement abnormal returns: A_1 is significant at the 5% level, and A_2 and A_3 at the 1% level. This is in line with theory and our findings on OMXSB inclusion pre-announcement abnormal returns. However, we do not find statistically significant results for any other variable than arbitrage risk here. One theory is that this could be, as previously noted, an issue of small sample size.

For OMXS30 exclusion pre-announcement abnormal returns, no findings are significant at the 5% level. Given that we only have 13 observations for OMXS30 exclusions, we believe this is also a problem of small sample size.

7.2.4. OMXS30 post-announcement abnormal returns

Unlike for OMXS30 post-announcement abnormal returns, we see that arbitrage risk has clear and statistically significant effects on post-announcement abnormal returns, for inclusions as well as exclusions and in both the (0, 15) and (15, 30) trading days relative AD event windows. In all cases except exclusions in the (15, 30) event window, A_2 and A_3 are significant at the 1% level. For exclusions in the (15, 30) event window A_1 is significant at the 5% level whereas A_2 and A_3 are not significant. Again, we believe the small sample size is an issue since this is in stark contrast to the theoretical prediction that A_1, A_2, and A_3 should be highly related. Due to these conflicting findings, we are unable to draw any conclusions about arbitrage risk in the (15, 30) event window for OMXS30 exclusions

We can also see that the interaction effect between arbitrage risk and demand shock is significant at the 1% level for OMXS30 inclusions in the (0, 15) event window. Amihud illiquidity, however, does not show any significant impact. However, as is also the case with A_1 through A_3 in both the (0, 15) and (15, 30) event windows for inclusions, we unexpectedly find that the model coefficients for arbitrage risk have negative signs. This is opposite to the price pressure hypothesis' theoretical prediction that liquidity providers should charge more for providing excess liquidity in high-arbitrage-risk stocks, which would hold that high-arbitrage-risk stocks should experience larger price gains in the (0, 15) event window but also correspondingly larger reversions in the (15, 30) event window. Instead, we observe negative coefficients in both the (0, 15) and (15, 30) event windows, meaning that OMXS30 stocks with high arbitrage risk experience larger price losses when included in the index. This is indeed a very perplexing finding since no existing theory predicts these effects, which we observe are significant at the 1% level.

For OMXS30 exclusions we observe the same negative coefficients as mentioned above, but in this case, they are in line with theoretical predictions since the demand shock is negative and arbitrage risk would be expected to amplify the negative demand shock effect.

Apart from the unexpected sign on arbitrage risk effect for OMXS30 inclusions, what is also strange to note is that for both inclusions and exclusions, arbitrage risk has clearly significant impact on OMXS30 post-announcement abnormal returns but not on OMXSB post-announcement abnormal returns. Theory provides no plausible explanation on why this would be the case, and it would seem that there are no material differences between the two indices that should give rise to this effect. Instead, we can hypothesize that this could be because OMXS30 inclusions may not experience the same type of price-pressure effects as OMXSB inclusions do in the 30 trading days that follow an inclusion announcement. Another plausible explanation would be that the effect we are observing might in fact be the reversion of a price pressure effect that could have started before the actual inclusion announcement. This could be possible since the mechanical nature of index revisions make them predictable at least to a limited extent. However, none of these two explanations would leave us any wiser as to why we observe the same relationship between arbitrage risk and post-announcement abnormal returns for inclusions as well as exclusions.

7.3. Suggestions for further research

To expand the current understanding of the connections between the index effect, arbitrage risk and transaction costs we have identified several areas in which further research could hold large potential. A closer analysis of the OMXS30 with a larger sample and better historical information would be of interest, and could hopefully provide insight on the seemingly-ambiguous results where arbitrage risk seems to be negatively related with returns in both our (0, 15) and (15, 30) trading days relative AD event windows.

It would also be interesting to look at what drives the strength of the relationship between arbitrage risk and abnormal returns: our models have vastly different explanatory power for the OMXSB and OMXS30 indices, even though the two indices are seemingly very similar.

Finally, it would be of great interest to investigate why explained variance influences post-announcement abnormal returns in the (15, 30) trading days relative AD window for OMXSB exclusions. This is a counterintuitive finding, given that explained variance can be easily hedged therefore should be of very limited relevance.

8. Conclusions

Our findings for the OMXSB as well as OMXS30 index show that arbitrage risk plays a large role in explaining the abnormal price again that a stock experiences during the 30 trading days leading up to its inclusion in either index. For OMXSB inclusions this effect on pre-announcement abnormal returns exists through interaction between arbitrage risk and demand shock size. Furthermore, we have shown that illiquidity also impacts the pre-announcement abnormal returns of OMXSB inclusions: as we predicted, illiquidity does limit arbitrageurs' willingness to engage in arbitrage trading that would flatten said stock's demand curve. A key reason that illiquidity does not show any impact for the OMXS30, we believe, is that illiquidity is much lower in the OMXS30 than in the OMXSB due to the index's nature.

In the case of exclusions from the OMXSB and OMXS30, our models have very limited explanatory power. We only find that the interaction effect between arbitrage risk and demand shock size influences pre-announcement abnormal returns for OMXSB exclusions, although arbitrage risk or demand shock size alone do not have any such effect.

We also find that for post-announcement abnormal returns the price pressure effects that OMXSB additions experience are not affected by arbitrage risk or transaction costs. This is not unexpected, given that both arbitrage risk as well as transaction costs in theory only have relationships with temporary and reverting price-pressure driven post-announcement abnormal returns – not any permanent price effects. However, for both OMXS30 inclusions and exclusions post-announcement abnormal returns have negative relationships with arbitrage risk. This is puzzling for two reasons: (i) we only see this effect for the OMXS30 and not the OMXSB, and (ii) while a negative relationship is predicted by theory for exclusions, a positive relationship would be predicted for inclusions.

Further, the negative relationship between arbitrage risk and post-announcement abnormal returns for OMXS30 inclusions leaves us bewildered because the theoretical expectation, is that the linkage between arbitrage risk and post-announcement abnormal returns should be positive. While we find several potential explanations for the positive linkage between arbitrage risk and post-announcement abnormal returns in the case of inclusions, we are unable to explain the existence of such a positive linkage for post-even abnormal returns following both inclusions and exclusions. For this reason, we are forced to conclude that the relationship of post-announcement abnormal returns with arbitrage risk and transaction costs is vague in theory and ambiguous in practice.

9. References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets Vol.5* (No.1), 31-56.

Amihud, Y., & Mendelson. H. (1986) Asset pricing and the bid-ask spread. *The Journal of Financial Economics, Vol. 17* (No.2), 223-249.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2005) Liquidity and Asset Prices. *Foundations and Trends in Finance, Vol. 1* (No 4), 269–364.

Bechmann, L. K. (2004). Price and Volume Effects Associated with Changes in the Danish Blue-Chip Index – the KFX Index. *Multinational Finance Journal, vol. 8* (No. 1 & 2), 3-34.

Beneish, M. D., & Whaley, R. E. (1996). An Anatomy of the "S & P Game": The Effects of Changing the Rules. *The Journal of Finance, Vol. 51* (No. 5), 1909-1930.

Brockman, P., Schutte, M.G., & Yu, W. (2009) Is Idiosyncratic Risk Priced? The international Evidence. Working Paper.

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), 1901-1930.

Choi, J. Y., Salandro, D., & Shastri, K. (1988). On the Estimation of Bid-Ask Spreads: Theory and Evidence. *The Journal of Financial and Quantitative Analysis, vol. 23,* (No.2), 219-230.

de Jong, F., Nijman, T., & Röell, A. (1995). A comparison of the cost of trading French shares on the Paris Bourse and on SEAQ. *International European Economic Review, vol. 39 (No.7),* 1277-1301.

Denis, D. K., McConell J. J., Ovtchinnikov. A. O., & Yu. Y., (2003). S&P 500 Index Additions and Earnings Expectations. *The Journal of Finance, Vol. 58* (No. 5), 1821-1840.

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

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

Evans, G. R., (2012). Stock Market Transaction Costs and Economic Growth. (PhD dissertation University of Georgia).

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, Vol. 25* (No. 2), 383-417.

Fraser, P., & Groenewold, N. (2000). The effect of exchange rate shocks on the volatility of Australian sector excess returns: a note. *Applied Economics Letters, 2000, vol. 7 (No. 2),* 77-81.

Harris, L. E., & 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), 815-829.

Hirshleifer, D. (2001). Investor Psychology and Asset Pricing. *The Journal of Finance, Vol. 56* (No. 4), 1533-1598.

Hurvich, C. M., & Tsai, C-L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika, Vol. 76* (No. 2), 297-307.

Jain, P. C., (1987). The Effect on Stock Price of Inclusion in or Exclusion from the S&P 500. *Financial Analysts Journal Vol. 43* (No. 1), 58-65.

Karlsson. J., Koria. R (2018). The Effect of Open-end Index Revisions A study of market efficiency on the Swedish Stock Exchange. (Bachelors thesis, Stockholm School of Economics).

Kaul, A., Mehrotra, V., & Morck, R. (2000). Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment. *The Journal of Finance, volume 55 (No. 2),* 893-912.

Kyle. A. S. (1985). Continuous auctions and insider training. *Econometrica Vol. 53* (No. 6*)*, 1315-1335.

Mackinlay, A. C. (1997). Event Studies in Economics and Finance. *The Journal of Economic Literature Vol. 35 (*No. 1), 13-39.

Mendenhall, R. (2004). Arbitrage Risk and Post-Earnings-Announcement Drift. *The Journal of Business, Vol. 77* (No. 4), 875-894.

Nasdaq. (2016). Rules for the Construction and Maintenance of the OMX Stockholm 30 Index. R*etrieved from: https://indexes.nasdaqomx.com/docs/Methodology_OMXS30.pdf*

Nasdaq. (2018). Rules for the Construction and Maintenance of the Nasdaq OMX All-Share, Benchmark and Sector Indexes. *Retrieved from: https://indexes.nasdaqomx.com/docs/Methodology_NORDIC.pdf*

Pontiff, J. (2005). Costly Arbitrage and the Myth of Idiosyncratic Risk. *The Journal of Accounting and Economics, Vol. 42* (No. 1-2), 35-52.

Roenfeldt. L. R., Gripentrog. G. L., & Pflaum. C. C. (1978). Further Evidence on the Stationarity of Beta Coefficients. *The Journal of Financial and Quantitative Analysis, Vol. 13* (No.1*),* 117-121.

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), 179-211.

Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *The Review of Financial Studies, Vol. 5* (No. 1), 1–33.

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

Shleifer, A., & Vishny, R. W. (1997). The Limits of Arbitrage. *Journal of Finance, Vol. 52* (No. 1), 35-55.

Stoll, H. R. (1989). Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests. *The Journal of Finance, Vol. 44* (No.1), 115-134.

Wurgler, J., & Zhurkavskaya, E. (2002). Does arbitrage Flatten Demand Curves for stocks. *The Journal of Business, Vol. 75* (No. 4), 583-608.

Appendix A – Regression of OMXSB exclusion pre-announcement Cumulative Abnormal Returns

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (-30, 0) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-15.887			-22.390		-27.326	-87.331		-17.451	
	(16.230)			(16.928)		(18.581)	(90.901)		(16.638)	
A_2		-15.762			-21.810					
		(16.367)			(16.992)					
A_3			-16.701							
			(16.818)							
Analyst dispersion							-0.087			
							(0.219)			
Demand shock size				-4.829	-4.707		4.367			
				(3.683)	(3.666)		(7.042)			
E_1						149.954				
						(119.423)				
A_1 * demand shock size								-11,055.740***		-11,047.010***
								(3,751.373)		(3,765.589)
Amihud illiquidity									88.594	40.968
									(194.144)	(184.653)
Constant	-0.057***	-0.057***	-0.057***	-0.029	-0.031	-0.069***	-0.054	-0.037**	-0.058***	-0.038**
	(0.017)	(0.017)	(0.017)	(0.027)	(0.027)	(0.020)	(0.071)	(0.016)	(0.017)	(0.017)
Observations	130	130	130	130	130	130	35	130	130	130
Adjusted R^2	-0.0003	-0.001	-0.0001	0.005	0.004	0.004	0.042	0.056	-0.007	0.049
Note:								**p<0.1; **p<0.05; ***p<0.01*		
								Standard errors in parentheses.		

Appendix B – Regression of OMXSB inclusion post-announcement Cumulative Abnormal Returns

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (0, 15) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	2.233			-8.014		15.838	-135.777**		7.528	
	(9.281)			(9.629)		(10.903)	(52.723)		(9.816)	
A_2		2.140			-8.156					
		(9.683)			(10.003)					
A_3			2.042							
			(9.822)							
Analyst dispersion							0.120			
							(0.175)			
Demand shock size				-5.764***	-5.734***		-6.953*			
				(1.895)	(1.887)		(3.744)			
E_1						-166.801**				
						(72.807)				
A_1 * demand shock size								-1,623.117		-2,069.007
								(2,049.109)		(2,059.688)
Amihud illiquidity									-158.235	-145.597
									(99.709)	(94.571)
Constant	0.028***	0.028***	0.028***	0.066***	0.066***	0.039***	0.106**	0.035***	0.028***	0.040***
	(0.011)	(0.011)	(0.011)	(0.016)	(0.016)	(0.011)	(0.043)	(0.010)	(0.011)	(0.010)
Observations	143	143	143	143	143	143	37	143	143	143
Adjusted R^2	-0.007	-0.007	-0.007	0.049	0.049	0.023	0.141	-0.003	0.004	0.007
Note:								**p<0.1; **p<0.05; ***p<0.01*		
								Standard errors in parentheses.		

	Dependent variable:									
	CAR (15, 30) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	0.538			0.177		-4.932	-70.206*		0.683	
	(7.949)			(8.515)		(9.474)	(37.820)		(8.482)	
A_2		0.880			0.558					
		(8.293)			(8.845)					
A_3			1.281							
			(8.412)							
Analyst dispersion							0.003			
							(0.126)			
Demand shock size				-0.203	-0.179		-0.701			
				(1.675)	(1.668)		(2.685)			
E_1						67.062				
						(63.263)				
A_1 * demand shock size								-1,771.933		-1,813.884
								(1,752.277)		(1,775.992)
Amihud illiquidity									-4.346	-13.698
									(86.164)	(81.545)
Constant	-0.029***	-0.029***	-0.030***	-0.028*	-0.028*	-0.033***	0.003	-0.023***	-0.029***	-0.023**
	(0.009)	(0.009)	(0.009)	(0.014)	(0.014)	(0.010)	(0.031)	(0.008)	(0.009)	(0.009)
Observations	143	143	143	143	143	143	37	143	143	143
Adjusted R^2	-0.007	-0.007	-0.007	-0.014	-0.014	-0.006	0.048	0.0002	-0.014	-0.007
Note:								**p<0.1; **p<0.05; ***p<0.01*		
								Standard errors in parentheses.		

Appendix C – Regression of OMXSB exclusion post-announcement Cumulative Abnormal Returns

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (0, 15) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-7.141			-4.740		-11.750	47.906		-4.609	
	(11.936)			(12.512)		(13.724)	(47.765)		(12.197)	
A_2		-6.833			-4.500					
		(12.036)			(12.555)					
A_3			-6.865							
			(12.372)							
Analyst dispersion							-0.158			
							(0.115)			
Demand shock size				1.783	1.816		1.866			
				(2.722)	(2.709)		(3.700)			
E_1						60.420				
						(88.205)				
A_1 * demand shock size								-2,653.082		-2,686.254
								(2,834.536)		(2,831.832)
Amihud illiquidity									-143.327	-155.787
									(142.325)	(138.864)
Constant	-0.044***	-0.045***	-0.045***	-0.054***	-0.055***	-0.049***	-0.052	-0.042***	-0.042***	-0.037***
	(0.013)	(0.012)	(0.012)	(0.020)	(0.020)	(0.015)	(0.037)	(0.012)	(0.013)	(0.013)
Observations	130	130	130	130	130	130	35	130	130	130
Adjusted R^2	-0.005	-0.005	-0.005	-0.010	-0.010	-0.009	-0.031	-0.001	-0.005	0.001

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.

	Dependent variable:									
	CAR (15, 30) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	10.175			6.457		31.362**	-28.658		8.680	
	(11.829)			(12.369)		(13.075)	(57.902)		(12.119)	
A_2		11.855			8.426					
		(11.915)			(12.402)					
A_3			12.190							
			(12.246)							
Analyst dispersion							0.166			
							(0.140)			
Demand shock size				-2.761	-2.668		3.046			
				(2.691)	(2.676)		(4.486)			
E_1						-277.729***				
						(84.036)				
A_1 * demand shock size								-900.528		-878.152
								(2,821.747)		(2,826.600)
Amihud illiquidity									84.662	105.086
									(141.410)	(138.608)
Constant	0.008	0.007	0.007	0.024	0.022	0.031**	-0.030	0.019	0.007	0.016
	(0.012)	(0.012)	(0.012)	(0.020)	(0.019)	(0.014)	(0.045)	(0.012)	(0.013)	(0.012)
Observations	130	130	130	130	130	130	35	130	130	130
Adjusted R^2	-0.002	-0.0001	-0.0001	-0.002	-0.0001	0.070	-0.023	-0.007	-0.007	-0.010

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.

Appendix D – Regression of OMXS30 exclusion pre-announcement Cumulative Abnormal Returns

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (-30, 0) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-99.005*			-111.739*		-133.561	425.950		-173.106**	
	(54.456)			(61.043)		(276.306)			(65.761)	
A_2		-137.264			-156.206					
		(83.712)			(94.668)					
A_3			-138.492							
			(84.343)							
Analyst dispersion							2.160			
Demand shock size				-16.335	-15.576		-21.976			
				(30.265)	(31.213)					
E_1						133.649				
						(1,045.614)				
A_1 * demand shock size								69,684.700		84,895.640
								(118,885.900)		(135,587.70
Amihud illiquidity									161,872.500	27,637.130
									(93,165.290)	(98,809.100)
Constant	0.079	0.084	0.083	0.126	0.131	0.076	-0.271	-0.052	-0.008	-0.087
	(0.094)	(0.100)	(0.100)	(0.130)	(0.140)	(0.101)		(0.105)	(0.100)	(0.166)
Observations	13	13	13	13	13	13	4	13	13	13
Adjusted R^2	0.161	0.123	0.124	0.103	0.059	0.079	0	-0.058	0.291	-0.155

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.

Appendix E - Regression of OMXS30 inclusion Cumulative Abnormal Returns for OMXS30 in (15, 30) trading days relative the announcement date

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (15, 30) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-59.232**			-59.736**		-75.545*	164.015		-50.146**	
	(21.262)			(21.752)		(37.392)	(221.266)		(21.681)	
A_2		-82.097***			-86.224***					
		(28.116)			(28.656)					
A_3			-82.605***							
			(28.413)							
Analyst dispersion							0.167			
							(0.534)			
Demand shock size				4.086	6.850		18.844			
				(7.806)	(7.659)		(23.712)			
E_1						126.568				
						(236.145)				
A_1 * demand shock size								-4,140.862		-5,978.991
								(4,455.816)		(4,074.262)
Amihud illiquidity									-57,555.730	-98,669.210**
									(41,115.850)	(43,407.220)
Constant	0.025	0.034	0.033	0.017	0.022	0.022	-0.143	-0.018	0.047	0.037
	(0.028)	(0.029)	(0.029)	(0.033)	(0.032)	(0.029)	(0.141)	(0.027)	(0.031)	(0.034)
Observations	19	19	19	19	19	19	7	19	19	19
Adjusted R^2	0.273	0.295	0.293	0.241	0.286	0.241	-0.280	-0.008	0.312	0.191

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.

Appendix F – Regression of OMXS30 exclusion post-announcement Cumulative Abnormal Returns

For variable explanations, please see the note to **Table 10**.

	Dependent variable:									
	CAR (0, 15) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-95.268***			-99.975***		-41.663	418.583		-63.280**	
	(22.872)			(25.722)		(114.844)			(27.383)	
A_2		-154.775***			-165.418***					
		(29.509)			(32.726)					
A_3			-156.122***							
			(29.674)							
Analyst dispersion							-1.145			
Demand shock size				-6.039	-8.751		-5.404			
				(12.753)	(10.790)					
E_1						-207.326				
						(434.600)				
A_1 * demand shock size								30,906.410		-47,136.020
								(70,770.720)		(52,721.380)
Amihud illiquidity									-69,878.550	-141,797.300***
									(38,794.800)	(38,420.550)
Constant	0.009	0.030	0.029	0.026	0.057	0.014	0.056	-0.100	0.047	0.079
	(0.039)	(0.035)	(0.035)	(0.055)	(0.048)	(0.042)		(0.062)	(0.042)	(0.064)
Observations	13	13	13	13	13	13	4	13	13	13
Adjusted R^2	0.577	0.688	0.690	0.545	0.678	0.545	0	-0.072	0.648	0.501

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.

	Dependent variable:									
	CAR (15, 30) trading days relative AD									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A_1	-60.211**			-60.839**		-78.284	-206.243		-97.399***	
	(23.783)			(27.041)		(120.631)			(27.296)	
A_2		-66.700			-63.751					
		(40.247)			(46.019)					
A_3			-69.496							
			(40.212)							
Analyst dispersion							0.176			
Demand shock size				-0.805	2.424		21.965			
				(13.407)	(15.173)					
E_1						69.897				
						(456.500)				
A_1 * demand shock size								41,597.820		44,159.240
								(56,798.370)		(64,999.270)
Amihud illiquidity									81,237.120*	4,653.896
									(38,671.060)	(47,368.030)
Constant	0.013	0.004	0.005	0.015	-0.004	0.012	-0.118	-0.066	-0.031	-0.072
	(0.041)	(0.048)	(0.048)	(0.058)	(0.068)	(0.044)		(0.050)	(0.041)	(0.079)
Observations	13	13	13	13	13	13	4	13	13	13
Adjusted R^2	0.311	0.127	0.142	0.242	0.042	0.244	0	-0.040	0.474	-0.143

Note: **p<0.1; **p<0.05; ***p<0.01*

Standard errors in parentheses.