EXCHANGE-TRADED FUNDS AND CONSTITUENT PRICE INFORMATIVENESS

FRIDA HABBE

SOFIA HELLENBORG

Bachelor Thesis

Stockholm School of Economics

2022

Exchange-Traded Funds and Constituent Price Informativeness

Abstract:

This thesis aims to investigate the relationship between exchange-traded funds and the price informativeness of their constituents using daily data on the proxies ETF coverage and price non-synchronicity over a period of 4.5 years, and finds some statistically significant evidence of a negative relationship. Additionally, it examines potential underlying drivers of this relationship using turnover and relative bid-ask spreads as dependent variables, finding weak and somewhat contradicting relationships. Thirdly, it expands from previous research by differentiating between domestic and foreign ETF coverage, and finds that different treatment of outliers yield statistically significant and somewhat differentiated implication of Swedish contra foreign ETF coverage for the price informativeness of Swedish stocks. Our findings are compared and contrasted to previous research on the ETF coverage-price informativeness relationship, and the merit of previously argued explanations for the relationship is evaluated.

Keywords:

Exchange traded funds, price informativeness, price non-synchronicity, information costs, trader behavior.

Authors:

Frida Habbe (24629)
Sofia Hellenborg (24636)

Tutor:

Riccardo Sabbatucci, Assistant Professor, Department of Finance

Examiner:

Adrien d'Avernas, Assistant Professor, Department of Finance

Acknowledgements:

We would like to express our gratitude towards our tutor Riccardo Sabbatucci for his valuable support and guidance during the process of researching and writing this thesis.

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

1. Introduction

Exchange-Traded Funds (ETFs) is an asset class of increasing popularity since its introduction in the 1990s (Ben-David et al, 2018) . Like mutual funds, the portfolio of an ETF consists of a number of risky assets. This pool of assets may be constructed to track an index, a specific industry, a commodity market, a currency, or any kind of strategy, similar to other types of funds. In contrast to other funds, ETFs trade on an open exchange like individual stocks, and can thus be traded directly between investor counterparties at any time during the opening hours of the exchange. (Avanza, 2022) Significant growth has persisted in recent years, with global ETF assets amounting to USD 7.6 trillion in 2021 (TrackInsight, 2022). In Sweden, ETFs have not yet accelerated to the same popularity as in the US and the UK (ETF Database, 2022), but are slowly increasing in demand from Swedish investors (Wallsten, 2022).

Liquidity is regarded as one of the most attractive features of ETFs compared to traditional funds and has proven to remain strong in spite of the pandemic-related volatility observed in 2020 (TrackInsight, 2022). The intraday liquidity offered is what sets ETFs apart from traditional index funds, and therefore may attract short-horizon liquidity traders. This view is supported by research investigating the duration of institutional investors' holdings which is significantly shorter for ETFs compared to the underlying securities (Ben-David et al, 2018 and Broman and Shum, 2018). Additionally, due to their often passive, index-tracking strategies, ETFs generally have significantly lower fees than mutual funds (Avanza, 2022), making them attractive to investors.

The interaction of buyers and sellers in financial markets allows the process of price discovery for the assets traded in that market. Since ETFs represent a basket of assets already traded individually on financial markets, conventional market theory would predict the price of the ETF to largely be determined by fluctuations in the prices of the individual constituent securities, with some degree of a lead-lag relationship. The idea that ETF ownership of a constituent should affect the price of the underlying asset is not consistent with the idea of a frictionless market in which the owner structure in itself should have little to do with the price of the underlying. However, the introduction of ETFs as an investment vehicle may influence the trading behaviors of market participants, and thus affect the pricing process of risky assets. Our measurement of interest when addressing a question of pricing effects is the price informativeness of the underlying stocks which represents the amount of firm-specific information that the stock price contains.

Israeli, Lee, and Sridharan (2017) propose the idea that uninformed investors, normally providing liquidity and acting as counterparties to informed traders in the underlying stocks, migrate to ETFs due to their attractive features such as diversification and ease of trade, creating larger trading costs in its constituents and therefore disincentivizing information gathering. Thus, the level of ETF coverage may deteriorate the pricing efficiency of the underlying assets.

Additionally, it is interesting to study whether ETFs domiciled in Sweden affect the informativeness of the Swedish underlying compound securities in a different way than international ETFs. Coval and Moskovitz broaden existing literature on the preference for domestic assets in international investment portfolios, shedding light on the fact that investors seem to prefer to invest "close to home" even in domestic investment portfolios. The study shows that investment managers' preferences suggest that there is asymmetric information between the local and foreign investors which may be the driver of said geographical investment preferences. It is on one hand possible that there is private information transmitted directly or indirectly between the parties, but it is also possible that investing locally means investors are more in tune with the information environment of the company. This could make it possible for local investors to deduce better estimates of stock value than foreign investors. (Coval and Moskovitz, 1999)

We therefore hypothesize that the effect of the migration of noise traders from the underlying securities to the ETFs would be most significant in the Swedish ETFs, as the local traders, both institutional and households, could be assumed to otherwise hold the compound securities prior to favoring the ETFs. Compare this to the foreign investors who would on average be more biased towards holding securities local to them.

The findings of our data analysis suggest that both Swedish and foreign ETF coverage, i.e. the ratio of a firm's shares outstanding that are owned through Swedish or foreign ETFs respectively, is negatively associated with the informativeness of that share price. Our initial regression fails to yield statistically significant results for the Swedish ETF coverage's effect on the price informativeness proxy price non-synchronicity, but is successful at the 5% significance level for the foreign ETF coverage. After winsorizing the price non-synchronicity variable the negative correlation is significant at the 10% significance level for the Swedish ETF coverage, and at the 5% significance level for the foreign ETF coverage. After both controlling the price non-synchronicity variable and the ETF coverage measurements for outliers we document an insignificant coefficient for the Foreign ETF variable, but a correlation between the Swedish ETF coverage and price non-synchronicity significant at the 5% level. From this, we conclude that some evidence of a negative relationship between domestic and foreign ETF coverage and price informativeness has been obtained.

Additional analysis is aimed at evaluating the aforementioned noise trader migration theory proposed by Israeli, Lee, and Sridharan (2017) by investigating the relationship between ETF coverage and factors thought to be related to information gathering costs, namely relative bid-ask spreads and share turnover. The additional regressions, while yielding statistically significant coefficients, indicate weak relationships and generate conflicting implications for the validity of the noise trader migration explanation, as Swedish ETF coverage proves to be related to higher bid-ask spreads (supporting the notion of information gathering disincentives) as well as higher turnover (contradicting the notion of information gathering disincentives). The opposite relationships are found for foreign ETF coverage on bid-ask spreads and turnover.

The contribution made by this paper to the body of research on pricing efficiency and ETFs is a geographical broadening and nuancing of existing knowledge. The first ETF to be introduced on a Swedish exchange was issued in the year 2000 by XACT, which still dominates the market for ETFs listed in Sweden (XACT, 2022). Due to the relative novelty of ETFs as an investment vehicle on Swedish exchanges, the existing research on this geographical market is limited. Secondly, by comparing ownership of Swedish listed companies through Swedish- versus foreign-listed ETFs, we make an attempt to nuance the prior findings by investigating a potential home-field aspect of ETF ownership on pricing efficiency. In line with our theoretical expectations, we find evidence indicating that foreign ETF coverage yields weaker negative implications for stock price non-synchronicity than Swedish ETF coverage.

The paper proceeds as follows: Section 2 consists of a literature review aiming to position this paper in the growing body of financial literature examining the relationship between ETFs and their underlying securities. Section 3 describes our methodology in terms of data collection and regression analysis, Section 4 provides an overview of the collected data, Section 5 summarizes our results, Section 6 discusses the interpretation of our study, its limitations, as well as potential areas of interest for future research, and Section 7 concludes our thesis.

2. Literature review

2.1 The Effect of ETFs on Constituents

Previous literature on the effects of institutional investors on underlying stocks has highlighted correlations. Ben-David, Franzoni, and Moussawi (2018) argue that the liquid nature of the ETFs attracts short-horizon liquidity traders, and therefore exposes the underlying to demand shocks which in turn increases the intraday and daily volatility stemming from arbitrage activity between the ETFs and the underlying. According to the work of Lin, Karim och Carter (2014) idiosyncratic return volatility has the potential to reflect stock price informativeness, which indicates that a possible effect of ETF coverage on the underlying stocks could be found.

Other authors highlight the ETF arbitrage mechanism as a contributing factor to stock price discovery for the stocks covered by the ETF. Chen and Strother (2008) examine the 50 largest stocks on the Shanghai stock exchange and their findings suggest that ETF ownership could improve the pricing efficiency of the underlying stocks, especially in the short run. Xu, Yin, and Zhao (2019) provide evidence that ETF trading, both one- and two-sided, speeds up price discovery, not only of the ETF itself but also positively impacts the index.

The results of the paper by Israeli, Lee, and Sridharan (2017) are somewhat contradicting when comparing relationships in different periods, but the main conclusion is that ETF coverage contributes to less informative pricing of the underlying assets. The text explores the connection between noise (uninformed) traders' migration to ETFs from the underlying and shows that increased ETF ownership can be linked to increased trading costs from broadened bid-ask spreads causing lower benefits from information acquisition for the underlying. This leads to less informative security prices for the ETF constituents. This is consistent with the noisy rational expectations examined by Grossman and Stiglitz (1980) and many others which predict that pricing efficiency is dependent on information costs. If uninformed traders move away from the underlying stocks in favor of the ETF, the noisy rational expectations would predict a decline in pricing efficiency.

2.2 Price Informativeness

The measure of price synchronicity, originally proposed by Roll (1988) as the R^2 of an asset pricing model, and later developed by Morck, Yeung, and Yu (2000) among others as a logarithmic modification of R^2, is a commonly used metric for measuring price informativeness. There are however some contrasting interpretations of the relationship between price synchronicity and price informativeness in the literature.

The most prominent school of thought argues that price non-synchronicity is the correct proxy for price informativeness, i.e. that price synchronicity and price informativeness have a negative relationship. Morck, Yeung, and Yu (2000) apply this interpretation to their findings that price synchronicity is higher in emerging capital

markets than in developed, reasoning that developed capital markets are better able to incorporate firm-specific information in stock prices due to political and institutional factors. Hutton, Marcus, and Tehranian (2009) further strengthen this interpretation with their findings that opacity in financial reporting is correlated to higher price synchronicity. Correspondingly, Zhou (2021) finds that an improvement in the information transparency of Chinese listed companies is associated with increased non-synchronicity. The intuitive basis of the idea of a negative relationship between synchronicity and price informativeness is that stock prices that move in close synchronicity with the market or industry are indicative of low contents of firm-specific information in the given stock's price.

Contrastingly, Dasgupta, Gan, and Gao (2010) demonstrate a positive relationship between information transparency (and thus stock price informativeness) and future stock price synchronicity. Using a quasi-natural experiment design, Kan and Gong (2018) similarly find that an improvement in the information environment of stocks is positively correlated with price synchronicity, arguing the notion of price synchronicity as the preferable proxy for price informativeness over price non-synchronicity.

While the findings indicating a positive relationship between informativeness and synchronicity are noteworthy, the vast multitude of empirical evidence strengthening the opposite case for a negative relationship renders this interpretation paradigmatic.

2.3 Positioning in the Existing Research Corpus

This thesis positions itself to the growing body of research on ETFs and their relationships with their underlying securities by joining Israeli, Lee, and Sridharan (2017) in its intersection with the literature on contributors and inhibitors of stock price informativeness. Furthermore, while Israeli, Lee, and Sridharan (2017) document a lagged deterioration of price informativeness associated with a previous increase in ETF ownership, our study documents a similar contemporaneous relationship. In addition, we provide geographical nuance to the existing literature and differentiate conclusions on subsets of investors, by studying the relationship between Swedish listed stocks and both Swedish and foreign domiciled ETFs. This is of interest as the Swedish ETF market is considerably younger and in an earlier stage of development than the US market which is frequently studied in the existing literature.

3. Methodology

3.1 Sample Collection

3.1.1 Swedish Sample

The Refinitiv database was used to identify ETFs based on the criteria that they are trading on Swedish exchanges, and are investing in equity listed on Swedish exchanges. Nine ETFs meeting these criteria were identified, all issued by the Swedish issuer XACT. Four of these were levered bull and bear versions of a fifth of the ETFs and were therefore excluded from our sample. The funds XACT Svenska Småbolag and XACT Norden Högutdelande were created in 2016 and 2017 respectively, and lacked constituent data in Refinitiv for large parts of the period and could thus not contribute conclusive enough data to be considered useful for our intended analysis.

The three ETFs remaining held a total of 183 stocks throughout the sample period. Excluding the stocks that were listed on non-Swedish exchanges held by the Nordics focused fund, and Swedish listed stocks where Refinitiv data for parts of the period was unavailable, we were left with a sample of 100 stocks. The daily weighted holdings in each stock and the daily market capitalization of our three ETFs throughout our sample period were collected. The data collected for each stock was daily closing prices, daily shares outstanding, daily market capitalization, IPO date, daily volume as well as last daily bid- and ask prices.

Regarding the time horizon of the data sample, we were restricted on one end by the fact that data on XACT ETF holdings is not available in Refinitiv before mid-June 2015, and on the other end by the fact that the latest Swedish House of Finance Fama and French factors dataset ends at the end of 2019. Given the relatively large number of underlying stocks, we were still able to achieve a sufficient number of firm-day observations.

3.1.2 Foreign Sample

To broaden our understanding of a potential home-field aspect of price informativeness for Swedish ETFs, our second step was to collect a corresponding non-Swedish sample of ETFs investing in our 100 Swedish stocks. The ETFs iShares MSCI Sweden ETF, iShares Global Timber & Forestry ETF, and WisdomTree Europe SmallCap Dividend Fund were selected from the ETF Database Country Exposure Tool based on their high Sweden exposure and constituent data availability in Refintiv. The daily weighted holdings in each of our sample stocks and the daily market capitalization of the three international ETFs throughout our time frame were collected.

3.2 Main Regression

To investigate the relationship between ETF ownership and stock price informativeness, our main analysis will regress the price informativeness proxy price non-synchronicity (PNS) against our main independent variables Swedish ETF coverage (SEC), and Foreign ETF coverage (FEC), and our control variables Age, Size, and BetaRisk. Our primary regression is defined in Equation 1.

$$PNS_{i,t} = \beta_0 + \beta_{SEC,i}SEC_t + \beta_{FEC,i}FEC_t + \beta_{AGE,i}Age_t + \beta_{Size,i}Size_t + \beta_{BetaVol,i}BetaRisk_t \quad (1)$$

3.3 Additional Regressions

To further nuance our understanding of the mechanisms behind the relationship between ETF coverage and price informativeness, we perform additional regressions with dependent variables thought to be connected to price informativeness. As argued by Israeli, Lee, and Sridharan (2017), the assumed migration of liquidity providing uninformed traders from the constituents to the ETFs raises information costs by reducing liquidity and increasing bid-ask spreads in the underlying. To examine the validity of this explanation for the ETF coverage and price informativeness relationship, we will perform the additional regressions specified in Equations 2 and 3, using relative bid-ask spread and turnover as dependent variables.

$$RBA_{i,t} = \beta_0 + \beta_{SEC,i}SEC_t + \beta_{FEC,i}FEC_t + \beta_{AGE,i}Age_t + \beta_{Size,i}Size_t + \beta_{BetaVol,i}BetaRisk_t \quad (2)$$

$$Turnover_{i,t} = \beta_0 + \beta_{SEC,i}SEC_t + \beta_{FEC,i}FEC_t + \beta_{AGE,i}Age_t + \beta_{Size,i}Size_t + \beta_{BetaVol,i}BetaRisk_t \quad (3)$$

3.4 Dependent Variables

3.4.1 PNS

As per the example set by Zhou (2021) and others, we have defined the dependent variable of interest used in our main regression (Equation 1), price non-synchronicity (PNS) as follows in Equation 4.

$$PNS_{i,t} = ln\left(\frac{1-R^2_{i,t}}{R^2_{i,t}}\right) \quad (4)$$

To estimate R^2 on each day of our sample, we perform a daily rolling Fama and French four-factor regression for each of our stocks, with a ten-day rolling window, defined in Equation 5.

$$R_{i,t} = \beta_0 + R_{f,t} + \beta_{MKT,i}MKT_t + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{MOM,i}MOM_t \quad (5)$$

The Swedish House of Finance, who provided the Fama and French data-set defines the independent variables as follows:

- R_f is the risk-free rate, for which 1-month Swedish treasury bills constitute the proxy.
- MKT is the Swedish market return, estimated by the return of the SIX Retrun Index.
- SMB is the average return on a value-weighted portfolio of small-cap stocks, subtracted by the average return on a value-weighted portfolio of large-cap stocks.
- HML is the average return on a value-weighted portfolio of value stocks, subtracted by the average return on a value-weighted portfolio of growth stocks.
- MOM is the average return on a value-weighted portfolio of high-return stocks, subtracted by the average return on a value-weighted portfolio of low-return stocks.

3.4.2 RBA and Turnover

The additional regressions specified in Equations 2 and 3 utilize the dependent variables relative bid-ask spread (RBA) and Turnover respectively. RBA is defined in Equation 6 as the difference between the last ask and bid price of the day for each stock relative to the last ask price of the day for the respective stock, and is an attempt at capturing the relationship between ETF coverage and trading costs. Turnover is defined in equation 7 as the portion of the total shares outstanding that are traded in a day for each respective stock, and is meant to represent stock liquidity.

$$RBA_{i,t} = \frac{Last\ Daily\ Ask\ Price_{i,t} - Last\ Daily\ Bid\ Price_{i,t}}{Last\ Daily\ Ask\ Price_{i,t}} \quad (6)$$

$$Turnover_{i,t} = \frac{Daily\ Volume_{i,t}}{\#\ Shares\ outstanding_{i,t}} \quad (7)$$

3.5 Independent Variables

3.5.1 ETF Coverage

A proxy for the number of constituent shares held by an ETF on any day was calculated by dividing the constituent's percentage of the ETF market capitalization by the closing price of that constituent stock. Building and expanding on the methodology of Israeli, Lee, and Sridharan (2017), we define two main independent variables of ETF coverage. In order to meaningfully interpret the coefficients obtained from the regression analysis and to normalize the variable distribution, we use a logarithmic definition of

these variables. Firstly, Swedish ETF coverage (SEC) is defined in Equation 8.

$$SEC_{i,t} = ln\left(1 + \frac{\#\ shares\ held\ by\ Swedish\ ETFs_{i,t}}{\#\ Shares\ outstanding_{i,t}} * 100\right) \quad (8)$$

Following the reasoning outlined in the introduction and detailed by Israeli, Lee, and Sridharan (2017), we expect that the SEC variable will display a negative relationship with PNS, i.e. that higher Swedish ETF coverage will be associated with lower stock price non-synchronicity and thus lower price informativeness. In regards to the additional regressions aimed at capturing the underlying drivers of this in terms of liquidity and information costs, we predict a negative relationship with turnover, and a positive one with relative bid-ask spreads.

To investigate a potential home-field aspect of price informativeness, we define our second main independent variable, Foreign ETF coverage (FEC) as follows in Equation 9. The FEC variable represents non-Swedish ETF ownership of each Swedish stock in our sample, as contrasted to the SEC variable representing Swedish ETF ownership over those same stocks.

$$FEC_{i,t} = ln\left(1 + \frac{\#\ shares\ held\ by\ non{-}Swedish\ ETFs_{i,t}}{\#\ Shares\ outstanding_{i,t}} * 100\right) \quad (9)$$

Again, in line with the argument made by Israeli, Lee, and Sridharan (2017) regarding the migration of uninformed, liquidity-providing traders from the underlying to the ETF, we expect the FEC variable to be negatively correlated with price non-synchronicity, as well as with and turnover. A positive correlation is expected to relative bid-ask spreads. However, in comparison to the SEC variable, we hypothesize that the FEC has weaker relationships with the dependent variables studied than the SEC. We reason that due to the well-documented home bias in investors' stock picking, foreign ETF coverage is less likely to cannibalize on the underlying, but rather attract investors that would otherwise not have invested in the underlying. Thus, the liquidity-trader migration from the underlying to the ETFs theorized to cause price informativeness deterioration, is expected to be less prevalent between Swedish stocks and foreign ETFs.

3.5.2 Control Variables

To control for the potential effect of company size, we define a control variable Size as the natural logarithm of the constituent market capitalization. In the context of investigating the relationship between ETF coverage and price informativeness, controlling for size is an attempt at capturing any effect of being included in a market capitalization-based index. Thus, when using this control variable, we can more confidently suggest attributing any documented relationship of ETF coverage to the ETFs as vehicles affecting investor behavior rather than the ETFs as reflections of existing indices based on market value. To control for the potential effect of firm age, we define a

control variable Age as the natural logarithm of listing years, calculated as days past since the constituent IPO divided by 365.25. To control for the effect of systematic risk on PNS, as recommended by Li et al. (2014), we define our third control variable, BetaRisk, to account for the firm's systematic risk. The BetaRisk variable consists of the daily market beta resulting from a rolling CAPM regression with a 10-day window.

4. Data

4.1 Data Overview

The sampled data, summarized in Table I, is a set of panel data covering 3 Swedish and 3 international ETFs and 100 of their Swedish-listed constituents. The first calendar day of the sample period is the 16th of June 2015, and the last is the 27th of December 2019. The number of available daily observations in the Swedish House of Finance Fama French data set within this time frame is 1,013, or 101,300 firm days of observations. Pandox and Nobina were both publicly listed on the 18th of June 2015, eliminating an additional 6 firm days of observations due to the logarithmic use of listing age. Thus, the regressions used in our analysis runs on 101,294 firm days of observations.

Table I: Sample Overview
The table presents statistics summarizing the data collected and analyzed in this thesis.
Firm day observations are the days in our sample period for which we were able to obtain both stock data from Refinitiv Eikon and data from the Fama French data-set from the Swedish House of Finance.

Metric	**Our data sample**
ETFs listed in Sweden	3
ETFs listed internationally	3
Constituent firms (share classes)	100
Sample period	2015.06.16 - 2019.12.27
Total firm days of observation	101,294
Panel data gathered for each ETF	Market capitalization, daily constituent weight.
Panel data gathered for each share class	Closing price, latest daily bid price, latest daily ask price, daily volume, shares outstanding, IPO date.

4.2 Descriptive Statistics

In Table II, the average PNS, SEC, FEC, Total ETF Coverage (TEC) defined as sum of SEC and FEC, Age, market capitalization, and market beta are listed on the constituent firm level. For interpretability, the ETF coverage variables SEC, FEC, and TEC are stated in terms of the percent of shares outstanding held by the respective subset of ETFs, disregarding the logarithmic transformation that is used in the main regression and specified in Equations 8 and 9. Regarding the PNS measurement, the cross-firm average

is -0.6396. Etrion has the highest value of 0.2391, indicating low synchronicity with the Swedish market and thus a high level of price informativeness. On the other end of the spectrum is Investor, with PNS-values of -2.0561 and -2.1696 respectively for the A and B share classes, indicating high synchronicity with the Swedish stock market and low price informativeness. A reasonable explanation for Investor's outlier status is the fact that it is an investment company, and thus an indirect investment in a larger portfolio of companies, many of which are part of our sample of constituent firms. Naturally, Investor's stock price does not contain high levels of firm-specific information, but should instead move in closer synchronicity with the market portfolio. Support for this explanation is found in the fact that Industrivärlden, another investment company, displays the third and fourth lowest PNS, -1.8577 and -1.6867 respectively for the A and C share classes.

The Average SEC throughout the period across the sampled constituents is 0.0524%. Swedbank shares are on average most owned by Swedish ETFs throughout the period at a 0.2588% average Swedish ETF ownership. Nokia is least owned by Swedish ETFs in the sample, with an average SEC of 0.0007% across the period. Foreign ETF coverage is largest in Bilia with a 0.3740% average. Combined, Svenska Cellulosa has the highest total ETF coverage of 0.5107% on average, while Nokia has the lowest average of 0.0007%.

Table II: Variable average per constituent firm

The table presents descriptive statistics for the 100 sampled constituents. PNS is the average price non-synchronicity for the period. ETF ownership is divided into three categories. Swedish ETF coverage (SEC), Foreign ETF coverage (FEC), and total ETF coverage (TEC). The measures SEC and FEC represent the average percentage of shares outstanding held by Swedish or Foreign ETFs, and the TEC representing the average ETF holding by both Swedish and foreign ETFs. Age represents the time in years since the initial public offering, and size the market cap of the company in MSEK. Beta is the market beta obtained from 10-day rolling CAPM regressions. The data is for the period 2015.06.16 - 2019.12.27.

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
AAK	-0.4003	0.0212	0.0000	0.0212	12	30,900	0.6645
ABB	-1.2525	0.0387	0.0000	0.0387	29	416,000	0.9575
AFRY	-0.2706	0.0009	0.1849	0.1858	32	14,100	0.7178
Autoliv	-0.7181	0.0726	0.0000	0.0726	23	79,600	1.0983
Africa Oil Corp	-0.3151	0.0274	0.0000	0.0274	7	5,050	1.0277
ASSA Abloy	-1.1216	0.2490	0.0718	0.3208	23	191,000	1.0075
Atlas Copco A shares	-1.5316	0.1703	0.0566	0.2268	44	327,000	1.4245

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
Atlas Copco B shares	-1.434	0.0381	0.0326	0.0707	28	327,000	1.4418
Atrium Ljungberg	-0.2715	0.0167	0.0408	0.0575	23	19,000	0.5105
Avanza	-0.2203	0.0088	0.2036	0.2123	25	11,300	0.8154
Axfood	-0.2081	0.0225	0.0000	0.0225	20	33,300	0.4416
AstraZeneca	-1.0617	0.0183	0.0000	0.0183	19	775,000	0.6648
Balder	-0.4941	0.0220	0.0000	0.0220	18	39,000	0.7303
Betsson	-0.3627	0.0371	0.3516	0.3886	17	10,300	0.9019
Bilia	-0.1549	0.0040	0.3700	0.3740	29	8,860	0.7200
Billerud Korsnäs	-0.883	0.0489	0.0000	0.0489	16	27,000	1.1473
Biogaia	0.0939	0.0191	0.0777	0.0968	19	5,520	0.4660
Biotage	-0.078	0.0087	0.0652	0.0739	17	4,530	0.8906
Boliden	-0.7214	0.1390	0.0914	0.2304	18	61,000	1.3117
Bure	-0.4539	0.0068	0.0000	0.0068	24	7,170	1.0343
Castellum	-0.5351	0.0573	0.0000	0.0573	20	36,500	0.6146
Cellavision	-0.0831	0.0083	0.0000	0.0083	10	3,890	0.8836
Clas Ohlson	-0.0621	0.0134	0.3344	0.3479	18	7,150	0.6087
Cloetta	-0.1558	0.0338	0.1332	0.1669	9	8,080	0.5696
Dustin	0.0164	0.0123	0.1760	0.1883	3	5,630	0.6182
Elekta	-0.3496	0.0517	0.0046	0.0563	24	33,100	1.0166
Electrolux	-0.6903	0.1278	0.0923	0.2200	36	72,400	0.9341
Etrion	0.2391	0.0062	0.0000	0.0062	7	615	0.4909

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
Evolution Gaming	-0.2423	0.0012	0.0602	0.0614	2	18,200	0.9164
Fabege	-0.4788	0.0376	0.0000	0.0376	28	31,600	0.6656
Fingerprint	-0.1995	0.0709	0.0000	0.0709	19	13,200	1.0446
Getinge	-0.8422	0.1080	0.0698	0.1779	24	34,700	1.0168
Gränges	-0.415	0.0280	0.2169	0.2449	3	6,460	1.0644
Hexagon	-1.1232	0.0128	0.0659	0.0787	29	138,000	1.1394
Haldex	-0.1762	0.0061	0.0616	0.0678	36	3,840	0.7480
Hennes & Mauritz	-0.997	0.1894	0.0472	0.2366	43	315,000	0.9333
Hansa Biopharma	-0.1092	0.0264	0.0000	0.0264	10	5,510	0.9968
Holmen	-0.6155	0.0252	0.3169	0.3422	81	30,900	0.9308
Hexpol	-0.8355	0.0438	0.0000	0.0438	9	27,800	1.0203
Hufvudstaden	-0.5477	0.0243	0.0000	0.0243	79	31,500	0.6621
Husqvarna	-0.674	0.0417	0.0704	0.1121	11	42,000	0.9803
ICA	-0.3088	0.0258	0.0398	0.0656	12	60,300	0.4515
Industrivärden A shares	-1.8577	0.0244	0.0000	0.0244	72	80,700	1.1548
Industrivärden C shares	-1.6867	0.0334	0.0377	0.0711	72	80,700	1.1348
Intrum	-0.6283	0.0387	0.0000	0.0387	15	27,500	1.1509
Investor A shares	-2.0561	0.0030	0.0000	0.0030	100	282,000	1.1208
Investor B shares	-2.1696	0.0433	0.0477	0.0910	100	282,000	1.1632
JM	-0.608	0.0574	0.1105	0.1680	36	16,200	0.9341
Karo Pharma	-0.0085	0.0131	0.0000	0.0131	19	3,870	0.6148
Kindred Group	-0.4355	0.0332	0.0000	0.0332	13	20,700	1.0223

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
Kinnevik	-1.1644	0.1017	0.1599	0.2616	20	69,500	1.1485
Kungsleden	-0.3226	0.0385	0.2636	0.3020	18	13,000	0.6337
Klövern	-0.2777	0.0153	0.0000	0.0153	20	15,200	0.7457
Latour Investment	-0.9167	0.0162	0.0000	0.0162	32	58,200	0.9995
Loomis	-0.4397	0.0119	0.1063	0.1182	9	21,000	0.9571
Lundin Mining	-0.2396	0.0014	0.0000	0.0014	14	31,400	1.2400
Lundbergföretagen	-1.3581	0.0184	0.0195	0.0380	34	43,300	0.9983
Lundin Energy	-0.5601	0.0646	0.0571	0.1217	16	68,300	1.0352
Mekonomen	-0.2356	0.0121	0.2691	0.2812	17	5,730	0.5134
NCC	-0.712	0.0295	0.0757	0.1052	29	21,600	0.8502
Nordea	-1.322	0.2368	0.0631	0.2999	20	362,000	1.0007
NetEnt	-0.4635	0.0247	0.2040	0.2287	10	12,000	1.0268
Nibe Industrier	-1.0444	0.0098	0.0000	0.0098	20	37,900	1.0822
Nobia	-0.2953	0.0143	0.2677	0.2821	15	13,100	0.9461
Nobina	0.0957	0.0143	0.3092	0.3235	2	4,490	0.6298
Nokia	-0.6598	0.0007	0.0000	0.0007	34	268,000	1.0553
Nolato	-0.22	0.0014	0.2472	0.2486	33	9,110	0.6978
Pandox	-0.1824	0.0140	0.0000	0.0140	2	12,900	0.6209
PEAB	-0.764	0.0184	0.2376	0.2560	30	20,000	1.0809
Ratos	-0.7854	0.0106	0.3429	0.3535	63	13,500	1.1381
RaySearch Laboratories	-0.0905	0.0104	0.0000	0.0104	27	3,620	0.7480
ReciPharm	-0.083	0.0170	0.0191	0.0361	3	5,840	0.7343

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
Saab	-0.7226	0.0151	0.0000	0.0151	19	36,900	0.9073
Sagax	-0.0328	0.0084	0.0000	0.0084	4	20,800	0.4728
Sandvik	-1.4845	0.2522	0.0719	0.3241	116	158,000	1.4776
SAS	-0.0477	0.0063	0.0000	0.0063	16	9,000	1.1697
Svenska Cellulosa	-0.7791	0.1784	0.3322	0.5107	67	113,000	0.9550
SEB	-1.5405	0.2255	0.0582	0.2837	45	202,000	1.0521
Securitas	-0.9669	0.1127	0.0798	0.1924	26	48,700	0.9544
Handelsbanken	-1.1729	0.2345	0.0675	0.3020	146	214,000	1.0204
Skanska	-1.0709	0.1149	0.0765	0.1915	52	71,900	1.0499
SKF	-1.2226	0.1203	0.0770	0.1973	103	74,100	1.4497
Skistar	0.0053	0.0020	0.1363	0.1383	23	6,520	0.4507
Swedish Orphan Biovitrum	-0.4592	0.0410	0.0000	0.0410	11	40,700	1.0993
SSAB A shares	-0.7699	0.0403	0.0000	0.0403	28	28,900	1.5377
SSAB B shares	-0.6985	0.0292	0.0000	0.0292	28	28,900	1.4585
Stora Enso	-0.7721	0.0058	0.0000	0.0058	19	87,200	1.2755
Sweco	-0.2929	0.0042	0.0208	0.0250	19	23,000	0.9237
Swedbank	-1.321	0.2588	0.0671	0.3258	22	213,000	0.9603
Swedish Match	-0.4109	0.0109	0.0951	0.1060	21	62,400	0.6176
Tele2	-0.7382	0.1068	0.0728	0.1797	21	54,400	0.8957
Telia	-0.8342	0.1988	0.0553	0.2541	17	174,000	0.7913
Thule Group	-0.3668	0.0243	0.1304	0.1547	3	16,500	0.8716
Millicom International Cellular	-0.4049	0.0394	0.0595	0.0989	13	48,300	0.7923

Constituent (share class)	PNS	SEC (%)	FEC (%)	TEC (%)	Age (y)	Size (MSEK)	Beta
Trelleborg	-1.3957	0.0488	0.0000	0.0488	53	40,700	1.2433
Vitrolife	-0.1456	0.0141	0.0042	0.0183	16	12,500	0.9860
Volvo A shares	-1.4051	0.0085	0.0000	0.0085	31	265,000	1.4312
Volvo B shares	-1.4963	0.2254	0.0652	0.2906	31	265,000	1.4353
Wallenstam	-0.4733	0.0258	0.0000	0.0258	33	23,900	0.6409
Wihlborgs Fastigheter	-0.4031	0.0337	0.2393	0.2730	12	15,600	0.6386
Total	-0.6396	0.0524	0.0778	0.1302	29	75,773	0.9302

Descriptive statistics on the aggregate level is detailed in Table III. Comparing the SEC and FEC value, we see a somewhat higher mean ETF coverage from foreign ETFs, paired with a lower median for FEC than SEC, indicating a more pronounced skew. Due to the limited prevalence of ETF coverage of Swedish stocks during the period, we also see a lower quartile of 0 for both ETF coverage variables.

Table III: Descriptive statistics aggregated by variable
The table presents descriptive statistics for the variables examined. PNS represents price non-synchronicity, and ETF ownership is divided into two categories; Swedish ETF coverage (SEC) and Foreign ETF coverage (FEC). The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers 100 Swedish stocks.

Variable	Mean	Std. Err	Std. Dev	Q.1	Median	Q.3
PNS	-0.6396	0.0040	1.2663	-1.4427	-0.6161	0.1929
SEC	0.0489	0.0002	0.0641	0.0000	0.0310	0.0539
FEC	0.0695	0.0003	0.1011	0.0000	0.0180	0.0923
Age	3.0074	0.0027	0.9098	2.6926	3.0472	3.4482
Size	24.1262	0.0042	1.3998	23.1183	24.0989	24.9843
BetaRisk	0.9302	0.0029	0.9148	0.5087	0.9333	1.3342

Table IV presents the correlations between the independent variables used in the main and additional regressions. The strongest correlations of 0.5291 and 0.4822 can be found between SEC and Size, and Age and Size respectively. Given the moderate strengths of these correlations, no concerns of collinearity in our sample are raised.

Table IV: Correlation Matrix of Independent Variables
The table shows a correlation matrix for our independent variables. The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers 100 Swedish stocks.

Correlation	SEC	FEC	Age	Size	BetaRisk
SEC	1.0000	-0.000	0.3106	0.5291	0.0933
FEC	-0.0000	1.0000	-0.0006	-0.1928	-0.0413
Age	0.3106	-0.0006	1.0000	0.4822	0.1080
Size	0.5291	-0.1928	0.4822	1.0000	0.1371
BetaRisk	0.0933	-0.0413	0.1080	0.1371	1.0000

5. Results

5.1 Main Regression

Overall, our regression model demonstrates a rather low explanatory value with an R^2 of 0.1778, meaning just over 82% of the PNS variance in the sample is explained by factors unaccounted for by the model.

However, the results of the regression analysis lined out in the methodology section show that ETF coverage does affect the PNS of the underlying stocks. Table V presents the results from our data where the dependent variable PNS is regressed against SEC, FEC, Age, Size, and BetaRisk. The table shows that the percentage of shares held by Swedish ETFs fails to prove statistically significant for the PNS of the underlying. Therefore we are not able to tell whether this coverage has implications for the PNS and if so if it implies positive or negative changes in the price non-synchronicity. These inconclusive findings are potentially due to the variability, sample size, or effect size.

The regression does however show that the foreign ETF coverage has a statistically significant negative coefficient of -0.0789 which indicates that a change in ETF coverage from the Swedish ETFs of one percentage point changes the PNS of the underlying stock by -0.0789 percent on average.

Following the logic of Morck, Yeung, and Yu (2000), a deterioration in PNS is an indicator of lowered price informativeness of the underlying, assuming that PNS is a correct proxy for price informativeness. Therefore, our results indicate that foreign ETF coverage negatively impacts price informativeness of the underlying stocks, but are inconclusive for the Swedish ETFs.

The lack of statistically significant findings for the SEC variable means we are unable to draw any conclusions regarding the population at this stage of the analysis. However, in the limited realm of our sample, the negative coefficient of -0.1082 is consistent with our expectations based on the reasoning of Israeli, Lee, and Sridharan (2017) that higher information costs due to noise trader migration from the underlying to the ETF lead to lower information content in the underlying stock price. The negative and statistically significant coefficient for the FEC variable indicates that the aforementioned relationship would be present for foreign ETF coverage. Due to the lack of comparability with a statistically significant SEC coefficient, a conclusion regarding the idea of a potential home-field aspect is difficult to draw at this stage.

The variables Age, Size, and BetaRisk make up our control variables, the former two of which are used in logarithmic form. The regression shows that a one percent change in firm listing age, holding all other variables constant, affects the PNS with a -0.1214 percent change. A percentage change in firm size affects the PNS with a -0.2120 percent change, and a 0.01 unit increase in BetaRisk affects the PNS with a -0.3698% change.

Table V: Regression Table: ETF ownership and controls on individual stock's PNS

The table reports estimates from the regression of ETF ownership on PNS. The frequency of the observations is daily and the sample covers the period June 2015 until December 2019. PNS represents the price non-synchronicity of the stocks, and ETF ownership is divided into the categories of Swedish ETF coverage (SEC) and Foreign ETF coverage (FEC). The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. The control variable Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers a sample of 100 Swedish stocks.

PNS	**Coefficient**	**Robust Std. Err**	**t**	**P>|t|**	**95% Confidence Interval**	
SEC	-0.1082	0.0682	-1.59	0.113	-0.2419	0.0254
FEC	-0.0789	0.0365	-2.16	0.030	-0.1504	-0.0075
Age	-0.1214	0.0045	-26.76	0.000	-0.1303	-0.1125
Size	-0.2120	0.0034	-62.08	0.000	-0.2187	-0.2053
BetaRisk	-0.3698	0.0082	-45.03	0.000	-0.3859	-0.3537
Constant	5.1928	0.0771	67.35	0.000	5.0417	5.3440
F(5, 101288)	**2932.98**					
Prob > F	**0.0000**					
R^2	**0.1778**					
Root MSE	**1.1482**					
n	**101,294**					

5.2 Additional Regressions

5.2.1 ETF Coverage and Stock Turnover

When regressing the stock turnover against our independent variables we are able to demonstrate a small R^2 of 0.0324, meaning that only 3,24% of the variation in turnover for the individual stock can be explained by this model. While the model does report low P-values, the low R^2 of the model means the conclusions drawn are not to be considered anywhere close to exhaustive in explaining turnover.

Table VI shows the results from this additional regression which indicates with statistical significance at the 1% level that both Swedish and Foreign ETF coverage carry very small implications for the stock turnover. The control variables Size and Age also prove to be significant at the 1% level of significance while BetaRisk proves insignificant. The table shows that the amount of shares held by Swedish ETFs has a small coefficient of 0.0153, meaning that a change in the Swedish ETF coverage by one percent is associated with an increase in the share turnover by 0.0153 percentage points. The foreign ETF coverage on the other hand displays an extremely small negative correlation

with the dependent variable, with a coefficient of -0.0022. This indicates that a one percent increase in Foreign ETF coverage implies a very modest decrease in turnover by 0.0022 percentage points.

The small positive regression coefficient for the SEC variable is not in line with the noise trader migration theory laid forth by Israeli, Lee, and Sridharan (2017) where traders move from the compound security in favor of the ETF decreasing the daily turnover of the individual stock. Instead, the results from this regression indicate that when Swedish ETF coverage increases, so does the stock turnover. On the other hand, the FEC variable shows a negative correlation with the activity, making it consistent with the noise trader migration theory. While the results are statistically significant, the low explanatory value of the model and the extremely small dimensions of the implied impact indicate that ETF coverage explains only a very small fraction of the variation in turnover.

The control variables consist of Age, Size, and BetaRisk, where the first two are used in logarithmic form. The regression table shows that a one percent change in firm age, holding all other variables constant, corresponds to a 0.0002 percentage point change in the turnover. A one percent change in firm size affects the turnover by -0.0009 percentage points, and the implications of BetaRisk are statistically indeterminate. Overall the effects of the control variables are very small.

Table VI: Regression Table: ETF ownership and controls on individual stock's Turnover

Table VI reports estimates from the regression of ETF ownership on Turnover. The dependent variable Turnover is defined as a ratio of daily trading volume divided by the number of outstanding shares for the individual security. ETF ownership is divided into the categories of Swedish ETF coverage (SEC) and Foreign ETF coverage (FEC). The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. The control variable Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers a sample of 100 Swedish stocks.

Turnover	**Coefficient**	**Robust Std. Err**	**t**	**P>|t|**	**95% Confidence Interval**	
SEC	0.0153	0.0004	43.45	0.000	0.0146	0.0160
FEC	-0.0022	0.0002	-11.39	0.000	-0.0026	-0.0018
Age	0.0002	0.0000	10.18	0.000	0.0002	0.0003
Size	-0.0009	0.0000	-38.19	0.000	-0.0010	-0.0009
BetaRisk	0.0001	0.0001	1.40	0.163	-0.0001	0.0003
Constant	0.0236	0.0005	43.72	0.000	0.0226	0.0247
F(5, 101288)	**510.19**					
Prob > F	**0.0000**					
R^2	**0.0324**					

Root MSE	**0.0058**
n	**101,294**

5.2.2 ETF Coverage and Relative Bid-Ask Spread

The second additional regression has a higher explanatory value with an R^2 of 0.3307, meaning the model is able to explain roughly one third of the variation of the dependent variable. This is the highest explanatory value achieved out of all the performed regressions, indicating that relative bid-ask spread is the one of our dependent variables that can most exhaustively be explained by ETF coverage and controls.

The SEC variable yields a 0.0008 coefficient against relative bid-ask spreads. This means that a one percent increase in SEC is associated with a very small 0.0008 percentage points increase in relative bid-ask spreads. The FEC variable yields a coefficient of -0.0026, meaning that a one percent increase in the foreign ETF coverage corresponds to a change of -0.0026 percentage points in the relative bid-ask spread.

Israeli, Lee, and Sridharan (2017) theorized that the migration of uninformed traders from the compound securities to the ETFs would increase the bid-ask spread of the underlying. As the coefficient for the SEC variable is positive, this does align with the noise trader migration belief, and as the opposite is true for the FEC variable, this could be consistent with the idea of home bias.

The control variables consist of Age, Size, and BetaRisk, where the first two of which are used in logarithmic form. The regression table shows that a one percent change in firm age, holding all other variables constant, corresponds to a -0.0003 percentage point change in the turnover. A one percent change in firm size is associated with a change in the turnover by -0.0009 percentage points. A 1 unit increase in BetaRisk would correspond to a disappearing small change in turnover. Overall the effects of the control variables are very small.

Table VII: Regression Table: ETF ownership and controls on individual stock's Relative Bid-Ask Spread (RBA)

Table VII reports estimates from the regression of ETF ownership on the relative Bid-Ask spread of the stock (RBA). The dependent variable RBA is defined as the day's latest bid-ask spread, divided by the latest ask price. ETF ownership is divided into the categories of Swedish ETF coverage (SEC) and Foreign ETF coverage (FEC). The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. The control variable Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers a sample of 100 Swedish stocks.

RBA	**Coefficient**	**Robust Std. Err**	**t**	**P>|t|**	**95% Confidence Interval**	
SEC	0.0008	0.0001	9.38	0.000	0.0007	0.0010
FEC	-0.0026	0.0001	-37.07	0.000	-0.0027	-0.0024
Age	-0.0003	0.0000	-22.15	0.000	-0.0003	-0.0003
Size	-0.0009	0.0000	-87.79	0.000	-0.0009	-0.0009
BetaRisk	-0.0000	0.0000	-8.19	0.000	-0.0001	-0.0001
Constant	0.0250	0.0002	102.27	0.000	0.0245	0.0255
F(5, 101288)	**5638.68**					
Prob > F	**0.0000**					
R^2	**0.3307**					
Root MSE	**0.0020**					
n	**101,294**					

5.3 Robustness Checks

5.3.1 Winsorizing PNS

To nuance our main results, we perform a 90% winsorization of our PNS variable, which limits the effect of outliers on our results. For this purpose, the new variable PNS_w is generated. The definition of this variable contrasts with the definition of the PNS variable in that the data in the bottom and top 5 percentiles are set to the 5th and 95th percentiles respectively. A comparison of the descriptive statistics of the PNS and PNS_w variables is detailed in Table VIII.

Table VIII: Descriptive statistics comparison
Table VIII reports on descriptive statistics of spread for the variables PNS and PNS_w. PNS represents the price non-synchronicity of the stocks, and PNS_w is a version of the PNS variable subjected to a 90% winsorization.

Variable	Mean	Std. Err	Std. Dev	Q.1	Median	Q.3	Kurtosis	Skewness
PNS	-0.6396	0.0040	1.2663	-1.4427	-0.6161	0.1929	3.9132	-0.1475
PNS_w	-0.6368	0.0035	1.1271	-1.4427	-0.6161	0.1929	2.2200	-0.0639

The main regression outlined in Equation 1 is repeated with the replacement of PNS by PNS_w. The updated regression is defined in Equation 10, and the results from this regression are presented in Table IX.

$$PNS\_w_{i,t} = \beta_0 + \beta_{SEC,i}SEC_t + \beta_{FEC,i}FEC_t + \beta_{AGE,i}Age_t + \beta_{Size,i}Size_t + \beta_{BetaVol,i}BetaRisk_t \quad (10)$$

For SEC, the regression coefficient is very slightly less negative at -0.1024 compared to -0.1082 for the main regression. Statistical significance at the 10% level is achieved with this treatment, with a p-value of 0.085. This documents a negative relationship between Swedish ETF coverage and price informativeness when noise from outliers in the dependent variable is neutralized.

For FEC, the winsorization results in a negative coefficient of -0.0807 compared to the previous -0.0789 meaning the new regression displays a marginally stronger relationship between foreign ETF coverage and price non-synchronicity with this treatment of outlier values for PNS. The p-value of 0.013 using our winsorized PNS shows statistical significance at the 5% level for this negative relationship. This increases our confidence that the documented relationship in the main regression is not driven by outliers.

This is the iteration of our main regression which yields significant coefficients for both SEC and FEC simultaneously, which allows for comparability. In congruence with our reasoning on lesser prevalence of noise investor base cannibalization by the foreign ETFs, this regression suggests that Swedish ETF coverage has stronger negative implications for constituent price informativeness than foreign ETF coverage.

The overall explanatory value of the regression model is similar but somewhat higher compared to the main regression.

Table IX: Regression Table: ETF ownership and controls on individual stock's PNS_w
The table reports estimates from the regression of ETF ownership on winsorized price non-synchronicity. PNS_w is a version of the price non-synchronicity variable subjected to a 90% winsorization. ETF ownership is divided into the categories of Swedish ETF coverage (SEC) and Foreign ETF coverage (FEC). The measures SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. The control variable Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers a sample of 100 Swedish stocks.

PNS_w	**Coefficient**	**Robust Std. Err**	**t**	**P>|t|**	**95% Confidence Interval**	
SEC	-0.1024	0.0595	-1.72	0.085	-0.2192	0.0143
FEC	-0.0807	0.0323	-2.50	0.013	-0.1441	-0.0173
Age	-0.1052	0.0040	-26.39	0.000	-0.1130	-0.0974
Size	-0.1930	0.0030	-63.68	0.000	-0.1990	-0.1871
BetaRisk	-0.3354	0.0074	-45.42	0.000	-0.3499	-0.3210
Constant	4.6567	0.0682	68.27	0.000	4.5230	4.7904
f(5, 101288)	**3147.68**					
Prob > F	**0.0000**					
R^2	**0.1833**					
Root MSE	**1.0186**					
n	**101,294**					

5.3.2 Winsorizing PNS, SEC and FEC

A second robustness check is performed where the SEC and FEC variables are subjected to winsorization in addition to the already winsorized PNS_w variable. The new variables SEC_w and FEC_w are generated by replacing the data in the bottom and top 5 percentiles with the 5th and 95th percentiles observations from the SEC and FEC variables respectively. Our main regression in Equation 1 is again adjusted to accommodate for the winsorized dependent and independent variables, as defined in Equation 11.

$$PNS\_w_{i,t} = \beta_0 + \beta_{SEC\_w,i} SEC\_w_t + \beta_{FEC\_w,i} FEC\_w_t + \beta_{AGE,i} Age_t + \beta_{Size,i} Size_t + \beta_{BetaVol,i} BetaRisk_t \quad (11)$$

The results are outlined in Table X. In this iteration of the regression, the SEC_w coefficient displays an even larger statistically significant relationship between Swedish ETF coverage and price non-synchronicity than the SEC variable in the previous regression. This gives the indication that exclusion of outliers in the dependent PNS or

independent SEC variables documents negative implications of Swedish ETF coverage for stock price non-synchronicity.

When it comes to the winsorized foreign ETF coverage variable FEC_w, its coefficient in the regression on the PNS_w variable is statistically insignificant at a p-value of 0.222. This is possibly due to the large number of firm day observations of 0 in the FEC variable, which causes the winsorization to remove much of the in-sample variability for FEC.

Table X: Regression Table: SEC_w, FEC_w, and controls on individual stock's PNS_w
Table X reports estimates from the regression of SEC_w and FEC_w on PNS_w. PNS_w, SEC_w, and FEC_w are versions of the price non-synchronicity (PNS), Swedish ETF coverage (SEC), and foreign ETF coverage(FEC) variables subjected to 90% winsorizations. SEC and FEC are defined as the natural logarithm of the gross percentage of shares outstanding held by Swedish or Foreign ETFs. The control variable Age is defined as the natural logarithm of the time in years since the initial public offering, and Size is defined as the natural logarithm of the market cap of the company. BetaRisk is the market beta derived from a 10-day rolling CAPM regression and estimates the systematic risk. The data is from the period 2015.06.16 - 2019.12.27 and covers a sample of 100 Swedish stocks.

PNS_w	**Coefficient**	**Robust Std. Err**	**t**	**P>|t|**	**95% Confidence Interval**	
SEC_w	-0.1607	0.0646	-2.49	0.013	-0.2873	-0.0340
FEC_w	-0.0442	0.0361	-1.22	0.222	-0.1151	0.0267
Age	-0.1055	0.0040	-26.52	0.000	-0.1133	-0.0977
Size	-0.1911	0.0031	-62.62	0.000	-0.1971	-0.1851
BetaRisk	-0.3353	0.0074	-45.41	0.000	-0.3497	-0.3208
Constant	4.6110	0.0688	67.05	0.000	4.4762	4.7458
f(5, 101288)	**3153.99**					
Prob > F	**0.0000**					
R^2	**0.1833**					
Root MSE	**1.0186**					
n	**101,294**					

6. Discussion

6.1 Interpretation of the Regression Analyses

As presented in the Results section and in line with our theoretical predictions, our regression analysis indicates a negative contemporaneous relationship between foreign ETF ownership and stock price non-synchronicity in the magnitude of approximately -0.08%, statistically significant at the 5% level. The interpretation of these results is that an increase in the portion of a Swedish share owned by Foreign ETFs is associated with a decrease in the firm-specific information content of that share price. At the first stage of our regression analysis, we fail to draw a relevant conclusion on the relationship between Swedish ETF coverage of Swedish stocks and price non-synchronicity, as the -0.11% coefficient found in our sample lacks statistical significance for the population. However, after subjecting the main dependent variable to a 90% winsorization in an effort to control for outliers, we are able to document a statistically significant negative relationship between Swedish ETF coverage and our price informativeness proxy, somewhat stronger than the foreign ETF coverage coefficient. While we cannot prove causality, this is in line with our reasoning that a lower than for demestic ETFs, but still prevalent noise trader migration from the underlying to the foreign ETF implies a decay of stock price informativeness.

Regarding our additional regression analysis aimed at investigating potential underlying drivers of stock price informativeness, we find contradicting and overall weak results on the implications carried by both Swedish and foreign ETF coverage. For Swedish ETF coverage, we find an association with very slightly larger relative bid-ask spreads. While the documented coefficient is extremely small, it is statistically significant at a p-value of 0.000. Keeping in mind that the practical relevance of this small coefficient can be questioned, this is in line with the explanation of noise trader migration increasing information costs in the constituents, and can thus serve as part of a potential explanation of the decline in price informativeness found some iterations of our main regression. However, we incongruously find a statistically significant but small positive coefficient for Swedish ETF coverage on share turnover in the magnitude of 0.02 percentage points. In contrast to our conclusions drawn from the reduced price non-synchronicity, the increased turnover would indicate improved facilitation of the price discovery process. When it comes to the role of foreign ETF coverage in the additional regressions, we find the complete opposite implications with a reduction in relative bid-ask spreads and turnover. While both coefficients are very small in magnitude, they blur the picture by juxtaposing the findings from Swedish ETF coverage, and by carrying opposing implications for pricing efficiency.

In sum, our main regression analysis is successful in documenting a significant negative relationship between foreign ETF coverage and price non-synchronicity in line with our theoretically based expectations. While our initial regression is unsuccessful in obtaining a statistically significant coefficient for Swedish ETF coverage on price non-

synchronicity, our following analysis using winsorized variables demonstrates a statistically significant negative relationship. However, the contradicting and very weak coefficients obtained in the secondary analysis of bid-ask spreads and turnover leave us unable to attribute the main results to the noise trader migration theory laid forth by Israeli, Lee, and Sridharan (2017).

6.2 Comparison to Previous Research

Similarly to Israeli, Lee, and Sridharan, we document a negative relationship between ETF coverage on stock price non-synchronicity. In terms of scale, Israeli, Lee, and Sridharan find a 9% decrease in non-synchronicity associated with a 1% increase in ETF coverage, while we find much smaller coefficients in our various regressions. However, the overall explanatory value of our non-synchronicity models are higher at R^2s of 0.1778 to 0.1833 compared to their adjusted R^2 of 0.086. Lastly, they find compelling evidence for the noise trader migration explanation, which we do not see reflected in our secondary regressions. (2017)

The most obvious explanations for our somewhat diverging results are firstly the fact that Israeli, Lee, and Sridharan (2017) study the time-lagged relationship, in contrast to the contemporaneous relationship studied in this thesis, and secondly the different developmental states and potentially dissimilar intrapopulation variations of the American ETF market studied by them, compared to the Swedish ETF market primarily studied by us.

6.3 Limitations and Suggestions for Future Research

6.3.1 Sample Limitations

There are some limitations to the construction of our sample to be mindful of when drawing conclusions from our results. Firstly, our sample is not randomized but rather composed on an availability basis. Due to the limited availability of ETFs listed on Swedish exchanges as well as of the reported data for these ETFs and their holdings, constructing our sample in this way was necessary to achieve a sample of a meaningful size. Thus, there may be unidentified factors affecting both the availability of data and the studied dependent variables. Similarly, there are some shorter periods of data gaps in the Swedish House of Finance Fama French dataset that may or may not be randomly occurring, thus potentially limiting the validity of our conclusions.

6.3.2 Lack of Evidence for Causality

While this study is sprung out of interest in the effects of ETFs on their constituents, the methodology of our study does not have the capability to prove a causal relationship between our studied variables. This means that while we achieve

some statistically significant results, they are not to be interpreted as an indication that it is ETF coverage, or the statistically significant control variables that actually cause the changes in price non-synchronicity, bid-ask spreads, or turnover. Instead, they should be interpreted as indications that correlations in certain directions between the independent and dependent variables are statistically likely to exist.

6.3.3 Low Explanatory Value

The generally low R^2 of all of our regression models (below 0.19 for all models except the model of relative bid-ask spreads) suggests that there are other, more important predictive factors of price informativeness than Swedish and foreign ETF coverage, firm age, market capitalization, and systematic risk. While our study does not aim at giving an exhaustive explanation of the facilitators and inhibitors of price informativeness, but rather study the relationship between ETFs on constituent price informativeness, it is essential not to overstate these relationships. However, as previously stated the R^2 of our price non-synchronicity regression models (0.1778, 0.1833, and 0.1833) exceed the adjusted R^2 of the Israeli, Lee, and Sridharan (2017) model of price synchronicity, indicating that R^2 values of smaller dimensions are to be expected in models studying the ETF and price (non-)synchronicity relationship.

6.3.4 Suggestions for Future Research

Since our various iterations of the main regression had mixed success in documenting statistically significant relationships between domestic and foreign ETF coverage and domestic stock price informativeness, it is still of interest to more accurately chart these relationships. A future study using a more extensive sample of foreign ETFs investing in a specific domicile could potentially define this relationship robust to winsorization and further investigate the validity of our hypothesis that this would yield similar but less accentuated results as domestic ETF coverage, due to existing but less emphasized cannibalization of ETFs on constituent noise investor bases. Future studies could also broaden the understanding of these relationships by investigating other geographical markets with different levels of ETF prevalence. Furthermore, with a longer period of data availability, more precise documentation of the timing of these and other implications could be studied to further shed light on how ETFs as investment vehicles interact with their constituents through investor behavior.

6.4 Importance of the ETF-Constituent Relationship

As ETFs grow in popularity as investment vehicles, it is of great importance to understand their implications for investor behavior and financial markets. While their ease of trading and generally low fees due to passive strategies make them attractive for retail investors, there may be adverse effects for these investors in terms of less well-priced products. On the other hand, if uninformed trader migration is the underlying

cause of the deteriorated price informativeness, it is arguably at the expense of retail investors that low ETF-covered stocks are more accurately priced since these investors are generally in a worse informational position than institutional investors. Nevertheless, increasing our understanding of how ETFs interact with their constituents is a question of importance for the democracy of financial markets since it likely carries
different implications for unequally powerful investor bases.

7. Conclusion

This thesis aimed at joining the growing body of research studying the relationships between exchange-traded funds and their constituent assets. The main relationship of interest was the one between the level of ETF ownership of a stock and the information content of that stock price, as measured by our proxies ETF coverage and price non-synchronicity, and the market scope was determined to be Swedish equity. Additionally, we set out to investigate the potential existence of a home-field effect by differentiating between Swedish and foreign ETF coverage of Swedish stocks. Thirdly, we attempted to shed further light on two potential drivers of price informativeness by investigating the relationships between ETF coverage and relative bid-ask spreads and turnover respectively.

The main achievement of our study was to successfully document negative relationships between Swedish and foreign ETF coverage and price non-synchronicity in different iterations of our regression. While our Swedish ETF coverage variable is insignificant before controlling for outliers, winsorized variable regressions document implications in the realm of -0.10% to -0.16% on price non-synchronicity per 1% increase in Swedish ETF coverage. Similarly, foreign ETF coverage proves associated with -0.08% implications for price non-synchronicity per 1% increase, before the coefficient becomes insignificant after reducing the in-sample variation through winsorization. Our interpretation of this is that increased ETF coverage tends to imply lower information content of the constituent stock price.

Our attempt to capture a home-field effect implying that foreign and Swedish ETF coverage would display different relationships with price informativeness was somewhat successful. When performing regressions with winsorized price non-synchronicity, we obtain statistically significant results where Swedish ETF coverage gives stronger negative implications than foreign ETF coverage. However, other iterations of the regression yields insignificant coefficients for SEC and FEC_w respectively. Thus, we will merely claim to have contributed a small piece in the mapping of this aspect of the ETF-price informativeness relationship, to be continued by future research.

When speculating about the potential reason for the negative relationship between ETF coverage and price informativeness, we have looked to Israeli, Ben, and Sridharan (2017) and their proposed noise trader migration. They find some evidence supporting the idea that uninformed traders who would otherwise be counterparties in informed trades migrate from the underlying to the ETFs, creating disincentives for information gathering in the form of increased trading costs, thus lowering the information content of the constituent share price. Since we find some evidence both in line with and opposed to this argument when studying bid-ask spreads and share turnover, we cannot make a confident assessment of its merit. Further research specifically targeting investor streams could nuance the findings made by Israeli, Ben, and Sridharan (2017).

As ETFs continue their popularization as diversified, easily traded low-fee investment vehicles, the fund-constituent relationship ought to be of increasing scientific interest. Many aspects of the nature, timing, geography, and implications for the democracy of financial markets of these relationships remain to be charted. Our study has contributed with documentation of a negative ETF coverage and price informativeness relationship in the developing Swedish equity ETF market.

References

Avanza, 2022. *Vad är och hur fungerar ETF:er/börshandlade fonder?* https://www.avanza.se/lar-dig-mer/avanza-akademin/borshandlade-produkter/vad-ar-hur-fungerar-borshandlade-fonder.html

Ben-David, I., Francesco, F., and Moussawi, R., 2018. Do ETFs Increase Volatility? *The Journal of Finance*, *73*(6), 2471–2535.

Chen, G., and Strother, T. S. 2008. On the contribution of index exchange traded funds to price discovery in the presence of price limits without short selling. Working paper: University of Otago, New Zealand.

Coval, J.D. and Moskowitz, T.J. 1999 Home Bias at Home Local Equity Preference in Domestic Portfolios. *Journal of Finance*, 54, 2045-2073.

Dasgupta, S., Gan, J., and Gao, N., 2010. Transparency, Price Informativeness, and Stock Return Synchronicity: Theory and Evidence. *Journal of Financial and Quantitative Analysis*, 45(5), 1189-1220.

ETF Database, 2022. *ETF Country Exposure Tool. https://etfdb.com/tool/etf-country-exposure-tool/*

Fang, Y., and Sang, G. C. 2012. *Index price discovery in the cash market.* Working paper: Louisiana State University.

Gong, S. and Kan, S., 2018. Does High Stock Return Synchronicity Indicate High or Low Price Informativeness? Evidence from a Regulatory Experiment, *International Review of Finance*, 18(4), 523-546.

Grossman, S., and Stiglitz, J. E. 1980. On the impossibility of informationally efficient markets. *American Economic Review*, 70, 393–408.

Hasbrouck, J., 2003. Intraday Price Formation in U.S. Equity Index Markets, *The Journal of Finance,* 58(6), 2375-2400.

Hutton, A., Marcus, A., and Tehranian, H., 2009. Opaque financial reports, R2, and crash risk. *Journal of Financial Economics,* 94(1), 67-86.

Israeli, D., Lee, C., and Sridharan, S., 2017. Is there a dark side to exchange traded funds? An information perspective, *Review of accounting studies,* 22(3), 1048-1083.

Lin, C., Massa, M., and Zhang, H., 2014. Mutual Funds and Information Diffusion: The Role of Country-Level Governance, *The Review of Financial Studies,* 27(11) 3343-3387.

Lin, K .J., Karim, K., and Carter, C. 2014. Stock Price Informativeness and Idiosyncratic Return Volatility in Emerging Markets: Evidence from China. *Review of Pacific Basin Financial Markets and Policies,* 17(4),

Morck, R., Yeung, B., and Yu, W., 2000. The Information Content of Stock Markets: Why do Emerging Markets have Synchronous Stock Price Movements? *Journal of Financial Economics*, 58(1), 215-260.

Roll, R., 1988, R^2. *The Journal of Finance*, 43(3), 541-566.

Track Insight. 2022. *Global ETF Survey 2021*. https://www.trackinsight.com/global-etf-survey-2021/ (Accessed: 2022-04-05)

Wallsten, B., 2022. Svenskar ratar nya spartrenden. *SVD Näringsliv*. January 8th.

Xu, L., Yin, X., and Zhao, J., 2019. Differently Motivated ETF Trading Activities and the Volatility of the Underlying Index, *Accounting & Finance,* 59(1), 859-886.

XACT, 2022. *Nordiska ETF-marknaden mars 2022*. Stockholm: XACT. https://assets.ctfassets.net/xygylbh89id0/7t69qGnYXepj8mokefWydG/75e8624571b23e6de04e3f62de634474/XACT_NYHETSBREV_mars.pdf

Zhou, H., 2021. Inquiry letter supervision and stock price informativeness-research based on the stock price non-synchronicity, *Applied Economics Letter*s, 2021, 1-6.