# Information Consumption and Price Informativeness

**THE LINK BETWEEN BEHAVIOR AND PRICE DISCOVERY**

**Renjo Gabro**

**Samuel Svensson**

Bachelor Thesis

Stockholm School of Economics

2021

**Information Consumption and Price Informativeness: The Link Between Behavior and Price Discovery**

Abstract:

As technologies to share and consume information become more efficient and the information available to investors grow exponentially, the relationship between information consumption and price discovery becomes increasingly important to understand. We study the relationship between information consumption and price informativeness. In particular, our paper seeks to investigate abnormal institutional attention and its relation to price non-synchronicity. We find that abnormal institutional attention affects price non-synchronicity positively. This effect is persistent over at least a week of trading. We also show that our results are robust to other proxies of price informativeness, as well as a placebo test. We find no economically significant effect on price non-synchronicity by retail investors. These findings are consistent with previous literature, while also pointing out the importance of considering investor behavior when studying price informativeness.

Keywords:

Price informativeness, information consumption, institutional investors

Authors:

Renjo Gabro, 24454
Samuel Svensson, 24517

Tutor:

Marieke Bos, Deputy Director, Swedish House of Finance

Examiner:

Adrien d'Avernas, Assistant Professor, Department of Finance, Stockholm School of Economics

Acknowledgements:

We are thankful to Marieke Bos for very helpful input and general guidance. We also want to extend our thanks to Bo Frantzén, Hugo Holmberg, and Magnus Fromin for helping us get the data we needed.

Bachelor Thesis

Bachelor Program in Business and Economics Stockholm School of Economics

# 1 Introduction

A financial market has many purposes, effects, and interactions with us, the agents within that market. One of its roles is to consolidate and convey information. This is done in many ways, and we do not yet fully understand them. One of the rather fundamental questions in finance is *how* information is incorporated into prices. This paper focuses on this topic, and more specifically on how and whether investors' information consumption affects stocks and their prices, as suggested by Ben-Rephael, Carlin, Da, and Israelsen (2021). This paper aims to extend this research by exploring how and whether information consumption affects *stock price informativeness*.

One of the instruments through which we gather information about our society and, more importantly, the real economy is through the prices of stocks (Bond, Edmans, Goldstein, 2012). The aggregate private and public information has a striking way of becoming incorporated into stock prices (Gyntelberg, Loretan, Subhanij, Chan, 2009; Fama, 1970). A noteworthy instance is found in the paper of Roll (1984), where it is demonstrated that private information regarding weather conditions by citrus-specific traders is incorporated in the financial market prices of citrus firms. This, then, improved the public's knowledge of weather conditions, showing that information can be accessed from financial market prices themselves. Another case is Dr. Alchian (1954)'s discovery of the fissile fuel used for the creation of hydrogen bombs by the U.S. in 1954. It was discovered by observing the stock prices, and inferred, even though the operations were highly secret. In fact, Alchian's original paper was confiscated and destroyed due to the sensitive nature of the subject (Newhard, 2014).

An interesting point of research is then the price informativeness of stocks (Chen, Goldstein, Jiang, 2007), and how investors behave when it comes to new information disclosed within the market, more specifically, the investor's information consumption and cross-learning behaviors (Ben-Rephael, Da, Israelsen, 2017).

Price informativeness is a concept that in its essence is a measure of how much information is contained within the price of a stock, implying that an informative stock price should reflect the fundamental value of that stock, and, thus, firm-specific information (Durnev, Morck, Yeung, and Zarowin, 2003). For this reason, within the scope of this paper, we define stock price informativeness as the amount of firm-specific information conveyed in the price of a given stock.

When new information surfaces the area, whether it is in the form of macro-economic announcements or a firm's earnings and forecast announcements, a Bayesian investor faces a "signal extraction problem". Bayesian investors are investors that renew their assigned probabilities of events - such as cash flows of an individual firm or related firms - based on new information. The signal extraction problem that the investors face, is then to determine to what extent the news announcement is firm-specific, and to what extent it is systemic, meaning it will affect other, related firms as well, and, effectively, their price informativeness (Savor and Wilson, 2016; Chan and Chan, 2014).

Given the background above, this paper builds upon two strands of literature: (1) the relationship between investor information consumption and stocks (Ben-Rephael et al., 2017, 2021) and, simultaneously, exploration of the possible venues of usage of ex post measures of investor attention; and (2) the drivers of price informativeness in stocks (Chen, Goldstein, Jiang, 2007). To contribute to the existing literature, this paper aspires to contribute to the bridging of the gap between these two strands of literature, by investigating the following research questions:

**HYPOTHESES:**

**Hypothesis I:** Individual firms' earnings announcements and non-earnings scheduled events, related firms' earnings announcements and non-earnings scheduled events, and macro announcements, are determinants of institutional attention.

**Hypothesis II:** Investors' information consumption of an individual stock will positively affect the price non-synchronicity of that stock, for institutional investors.

**Hypothesis III:** Investors' information consumption of an individual stock will positively affect the price non-synchronicity of that stock, for retail investors.

In order to investigate the research questions provided above, this paper will firstly conduct Logit panel regressions to confirm whether information consumption, in conjunction with news announcements released by firms or institutions, is a suitable proxy for institutional investor attention and investors' cross-learning, as suggested by Ben-Rephael et al. (2017, 2021). A detailed description of how this test is constructed, will be provided in the upcoming section "Data and Sample Construction" under *Abnormal Institutional Attention* (AIA).

Further, to examine whether price informativeness is positively influenced by the information consumption of a stock, this paper uses *price non-synchronicity* as a proxy for price informativeness, as previously demonstrated by Chen, Goldstein, Jiang (2007). Price non-synchronicity has wide theoretical foundations as a measure of private information in a stock's price. In accordance with Kan and Gong (2018) and Ferreira, Ferreira, and Raposo (2011), this paper uses the Fama-French three factor model to compute price non-synchronicity.

To determine whether there is an influence on price informativeness by information consumption, this paper uses similar regressions specifications as Ben-Rephael et al. (2021), with the exception of exchanging stock price returns with price non-synchronicity as the dependent variable. The usage of price non-synchronicity as a dependent variable is motivated through well-adopted use in previous research papers, such as Kan and Gong (2018). To assist the interpretations of our findings, this paper includes a placebo test as a robustness check, to see if our results are robust or noise driven. We also include collinearity checks and double clustered standard errors to strengthen our choices of regression variables.

The results of this study suggest a positive correlation between information consumption, through the proxy *abnormal institutional attention*, and price informativeness, through the proxy *price non-synchronicity*. Additionally, the results suggest a persistence of the effect of institutional attention on price non-synchronicity, over a period of time, at least a trading week. The results also indicate that earnings announcements and non-earnings scheduled events are determinants of institutional attention as well as related firms' news announcements, which is in line with Ben-Rephael et al. (2021). However, the results do not indicate macro announcement as a determinant of institutional attention, which is not in line with Ben-Rephael et al. (2021). We find that for retail investors, there is no economically significant effect on price non-synchronicity.

Our findings provide a new context to the novel usage of abnormal institutional attention as a proxy to institutional information consumption, as well as a foundation for further research on the relationship between investor information consumption and price informativeness, or, in more general terms, the relationship between investor behavior and price discovery.

## 1.1 Literary Review

The main strands of literature in this paper are, as previously mentioned, that of the relationship between investor information consumption and stocks—with a basis on Ben-Rephael et al. (2021)—and that of price informativeness. A literary review is presented below, as a necessity and useful background to gain understanding of how this paper aspires to contribute to the literature.

### I. Information consumption

Historically, a multitude of papers have covered investor information consumption in many different ways, and with different proxies. Sanders and Zdanowicz (1992) were early with using abnormal returns, trading volumes and *Wall Street Journal* stories as a proxy for investor attention. Barber and Odean (2008) developed this method, using abnormal trading volume, previous day's returns and news coverage as a proxy for attention. Barber and Odean (2008) found evidence for higher trading volumes for stocks of firms with higher news coverage, indicating investor attention also affects prices. Engelberg and Parsons (2011) and Engelberg, Sasseville, and Williams (2011) found that investor attention, and trading volume, increases depending on the extent of a firm's coverage in local newspapers as well as on television shows, respectively.

Ben-Rephael, Da, and Israelsen (2017) presented a new, ground-breaking contribution to the literature of investor attention, with a measure of *institutional* investor attention called *Abnormal Institutional Attention* (AIA). In their paper, they propose AIA has advantages as a proxy of institutional investor attention, intuitive with the characteristics of institutional investors. The reason for this is that AIA is constructed using Bloomberg data on activity in its terminals. These terminals are expensive and have an expected annual cost of $20,000-$24,000 per machine, meaning they are likely to be used by institutional investors.

Da, Engelberg, and Gao (2011) presented a new, modern-day-adjusted measure as a proxy for *retail* investor attention, namely *Google Search* frequency. This proxy has been subject to numerous studies, investigating its legitimacy. Chae, Kim, and Han (2020) developed this measure, by using Google Search frequency and confining it to trading volume by individual investors. In Ben-Rephael et al. (2021), Google Search Volume Index (DSVI) is used, and further developed to match and gain the ability to be compared with AIA, by constructing it in a similar way, turning it into a dummy variable for each day depending on its relative volume compared with the previous 30-trading-day average volume.

The distinction between retail and institutional investors is important, and based on the recent literature, also measurable. Retail investors are characterized in different ways, with different resources and incentives. Retail investors are also likely to be more limited by time and attention constraints (Liu and Peng, 2015). Further, Ben-Rephael et. al (2017) finds that, on major news events, institutional investors lead retail attention, but not vice versa. This has implications in the field of information consumption, encouraging the distinction of institutional and retail investors.

### II. Price informativeness

Stock price informativeness as a concept is well-studied, while there remains debate on the proxies for measuring price informativeness. Durnev, Morck, Yeung, and Zarowin (2003) suggested, early, that price informativeness should be a measure of the fundamental value of a stock, and, thus, firm-specific information.

Gelb and Zarowin (2002) laid the modern groundwork of studying price informativeness by presenting a definition for price informativeness as: "the association between current stock returns and future earnings changes: more informative stock price changes contain more information about future earnings changes".

A widely used proxy to measure price informativeness is *price synchronicity* or, respectively, *price non-synchronicity*. Roll (1988) suggested that the degree of which stocks' returns are positively correlated, is related to amounts of market-specific or firm-specific information reflected within the price of a stock. Notably, Roll admitted this finding lacked substance, due to the possibility of it being a consequence of noise, rather than the amount of firm-specific information.

Morck, Yeung, and Yu (2000), Wurgler (2000), Durnev, Morck, Yeung, Zarowin (2003), Durnev et al. (2004), and An and Zhang (2013) pioneered the use of price synchronicity as a proxy for price informativeness, by providing evidence that price synchronicity relationship with stock returns were, in fact, correlated, and not merely the product of noise trading, and, thus, building their various studies upon this finding.

Jan and Myers (2006) and Dasgupta, Gan, and Gao (2010), on the contrary, suggested that price synchronicity is rather negatively correlated with price informativeness. Theoretically, when operating in efficient markets, where transparency is ubiquitous, then more firm-specific information is already available to the market. Dasgupta et al. (2010) argue that, in this case, firms with more transparent environments should have more informative stocks, and have a *higher* price synchronicity, since the market has already factored in much of firm-specific information, and should not be "shocked" by earnings announcements. This view is further developed by, amongst others, Hou, Peng, and Xiong (2013), and Kan and Gong (2018) which further cast doubt on the common view that price synchronicity is negatively correlated with price informativeness.

## 1.2 Structure of paper

The remainder of this paper is structured as follows: in section 2, we explain our ideal experiment, our data, variables, and sample construction; in section 3, we further develop our research questions, and discuss the methodology to answer these questions; in section 4, we present the results of our main analyses and provide a discussion of their implications; in section 5, we present robustness checks conducted to support our results; in section 6, we conclude; and in section 7, we give an account of all references used in this paper.

# 2. Data, Sample Construction, and Summary Statistics

In this section, we provide development of our data, sample, and variable construction, as well as summary statistics. The section starts with a description of an ideal experiment, followed by an account of the data sources and sample construction, and, finally, a presentation of the summary statistics.

## 2.1 Ideal Experiment

There are multiple factors that impact our results and the interpretation thereof, due to a deviation from an ideal dataset, as a consequence of a lack of access to said dataset. We investigate the major implications below.

**I. Omitted variables**

Bhushan (1989) and Gompers and Metrick (2001) argued that a set of firms might have certain characteristics which influences the ability of an investor, or analyst, to interpret new information regarding them. An ideal experiment, to investigate information consumption of institutional investors, would then include making previously unknown information for a *random* set of firms, accessible to investors, and learn from their trading behavior. Further, the experiment would include a control group of investors trading without the released information, and learn from their trading activity.

As such an experiment is out of the scope of this paper, we are humble of this weakness in our results, and understand that the results might be explained to some extent by omitted variables. To negate some of this bias, we introduce a fixed effects analysis, which makes these variables constants.

**II. Lack of access to PIN data**

Due to the criticism of price synchronicity as a proxy for information consumption, introduced by Jan and Myers (2006), Dasgupta et al. (2010), Hou et al. (2013), and Kan and Gong (2018), an ideal experiment would resort to an additional measure of price informativeness as a comparison and robustness check. One such measure, with strong theoretical foundations and wide use, is *Probability of Informed Trading* (*PIN*), which is regarded to convey the portion of private information reflected in the price of a stock.

In this paper, however, we lack access to the NYSE Trade and Quote (TAQ) database, which renders us unable to include this analysis. Therefore, we remain humble of an implied weakness in our results and conclusions, due to a lack of this robustness check. To account for some of this weakness, we introduce a placebo test to check if our results are driven by noise. Further, we introduce an alternative measure of price informativeness, Amihud (2002)'s illiquidity ratio. This lets us triangulate our results and produce more robust interpretations. The usage of Amihud's illiquidity ratio as a robustness check is motivated through wide usage in prior literature (Kelly, 2014; Li, Rajgopal, and Venkatachalam, 2014; Kan and Gong, 2018; Gassen, Skaife, Veenman, 2020).

### III. Google Search Volume Data

An ideal experiment would include daily search volume data for all tickers of stocks included within our experiment for the entire time period, in accordance with Ben-Rephael et al. (2017, 2021). Due to complications in Google Trends API scraping, only daily data for the period 2012-12-31 to 2017-12-31 was obtained for this paper, which means that we are missing data from more than two years compared to our main dataset. This is a deviation from the dataset of Ben-Rephael et al. (2021).

This has significant potential implications on this paper's ability to answer how retail investors' information consumption affects the price non-synchronicity of stocks. Our foundations for testing Hypothesis III will, therefore, be weakened. We are humble for this lack of robustness, and emphasize that our results and conclusions, especially for Hypothesis III, should be interpreted as indications and a point of interest for further study, rather than truths about the financial markets.

### IV. Other factors

Other factors affecting our results include the issue of *reverse causality*. We understand that previous price informativeness of a firm might influence the information consumption by investors. Due to difficulties in investigating this problem, the implication on our results is a weakness in the identification of the drivers of price informativeness.

As this paper largely follows the methodology of Ben-Rephael et al. (2021), an ideal experiment would have the exact same sources of information. However, our firm data have small deviations, even after following the same time intervals and being downloaded from the same source - Bloomberg - and following the same criteria. This might be attributed to being downloaded from Bloomberg at different times, which might have influenced the data. Further, this paper lacks access to news coverage data of RavenPack. We remain humble that this implies a weakness in our regression's coefficients due to the problem of omitted variables.

## 2.2 Sample Construction and Data Sources

Following the dataset of Ben-Rephael et al. (2021), this paper bases its analyses on the firms included in the Russell 3000 Index in the Bloomberg terminal, for the period Feb 2010 - Dec 2017. This time period corresponds to that of Ben-Rephael et al. (2021). The sample is delimited by data availability for the AIA variable - AIA data is only available from the 17th of February 2010, and there are a few periods in-between the timeframe, in which some data is missing.

Further, in correspondence with Ben-Rephael et al. (2021), this paper obtains data on stock prices, trading volumes, shares outstanding, and the like from the CRSP database. Fundamentals such as assets, book values, debt are obtained from the Compustat database. Data of information events, such as earnings announcement days, is gathered from I/B/E/S, while data on macroeconomic announcements is gathered from Bloomberg.

Following Ben-Rephael et al. (2021) and Da, Engelberg, and Gao (2011), our sample includes stocks that satisfy the following criteria: 1) have nonmissing values on Bloomberg terminals and Google Search Volume Index; 2) have a share code of 10 or 11 in the CRSP database; 3) have a stock price which is greater or equal to \$5 at the end of the previous month; and 4) have nonmissing book-to-market values. After applying these criteria, our sample includes 2,464 unique stocks and 3,624,621 observations.

Since the dataset is downloaded several months apart in this paper and Ben-Rephael et al. (2021), there is a small risk of unaccounted changes in Bloomberg terminals (as well as in CRSP, Compustat, and I/B/E/S). This implies a small inconsistency in the results. However, due to the large number of observations, this risk is significantly reduced.

For a complete list of all variables used in our regressions, see Appendix A.

## I. Abnormal Institutional Attention (AIA)

Our main independent variable used in the analysis is the measure for abnormal institutional attention, AIA. The sample construction for this variable follows Ben-Rephael et al. (2021). The variable captures investor attention ex post by using data from Bloomberg.

Bloomberg themselves construct a measure of attention by recording the number of times news for a specific firm is read and researched. In Bloomberg, one can read news from a specific firm by searching for the firm combined with the function "CN", for "Company News". However, readers can also read an article without explicitly searching for a specific firm. To differentiate between these two situations, Bloomberg assigns a score of ten when a user explicitly searches for firm-specific news, and they assign a score of one when the user just reads a news article.

Bloomberg then aggregates these scores into hourly counts. Using these hourly scores, Bloomberg creates a numerical attention measure by comparing the average hourly count for the last eight hours with the hourly counts for the past 30 days for that specific stock. This eight-hour rolling average is then assigned a score of 0 if the average is in the lowest 80% of all hourly counts the last 30 days. In a similar fashion, if the rolling average is between 80% and 90%, 90% and 94%, 94% and 96%, or larger than 96% of the previous 30 days' average, then Bloomberg assigns it a score of 1, 2, 3, or 4, respectively. Then Bloomberg takes a maximum of all scores from the calendar day, to aggregate the data up to a daily frequency. Bloomberg does not provide the hourly scores or the raw data, they only provide the aggregated data.

From the data provided by Bloomberg we follow Ben-Rephael, Da, & Israelsen (2017, 2021) and create a dummy variable *AIA* that is equal to one when the Bloomberg data is 3 or 4 and equal to zero otherwise. In other words, the dummy variable *AIA* is equal to one when institutional investor attention during one day is greater than 94% of all hourly counts in the past 30 days. This ensures that the focus of this paper is on days when investor attention is abnormal.

Since Bloomberg is a private firm, and does not provide information about their user base, we stay humble for concerns regarding the veracity of their measure of institutional attention. However, Ben-Rephael et al. (2017), conduct an extensive review of Bloomberg user profiles. They find that over 80% of the users work in financial industries, around 7% work in tech (of which, almost 80% are Bloomberg employees), and 1% of the users have academic email addresses linked to their profiles. These findings indicate that the Bloomberg users are likely to be, or work for, institutional investors that have the incentives and resources to react to news and incorporate information into prices. Furthermore, since Bloomberg terminals are expensive and have an expected annual cost of $20,000-$24,000 per machine, this too indicates a user base mainly of resourceful, institutional investors. With this research in mind, we find it substantiated that the measure of institutional attention is attributed to institutional investors, and, therefore, that the AIA variable measures what we intend it to—abnormal institutional investor attention.

## II. Abnormal Retail Attention (ARA)

The second main independent variable, used as a control variable in this paper, is *Abnormal Retail Attention* (ARA). The construction of ARA is derived from Ben-Rephael et al. (2021) and Da, Engelberg, and Gao (2011). Its data source is Daily Google Search Volume Index, and measures ex post spikes in retail attention.

Following the framework provided by Da et al. (2011), we define the *Abnormal Daily Google Search Volume Index* variable as follows:

$$ADSVI_{i,t} = \ln\left[\frac{DSVI_{i,t}}{monthly\ average\ DSVI_i}\right] \qquad \text{(Eq. 1)}$$

To ensure the ability of ARA to be compared to AIA, we follow the methodology provided by Ben-Rephael et al. (2021), and first assign a score of 0, 1, 2, 3, or 4 to ADSVI, according to the same rules used in the construction of the AIA measure; rolling average intervals of the previous 30 days' average as a base: below 80%, 80% and 90%, 90% and 94%, 94% and 96%, or larger than 96%, respectively. We then compute the dummy variable *ARA*, so that it has a value of zero when ADSVI has an assigned score of 0, 1, or 2; and a value of one when ADSVI has an assigned score of 3 or 4. As per the reasoning for AIA, this ensures that ARA encompasses days when the ADSVI is greater than 94% of the previous month's values, and can, thus, be regarded as abnormal retail attention.

## III. Information Events

Abnormal attention, from both retail and institutional investors, can arise in many contexts. These contexts include rumors or important news about firm-specific events, firm announcements, or macroeconomic announcements. Due to the variety of contexts for abnormal investor attention, we follow Ben-Rephael et al. (2021), and construct three proxies for events that are believed to influence abnormal investor attention: (1) firm earnings announcements; (2) other firm non-earnings scheduled events; and (3) macroeconomic announcements.

Earnings announcement dates are retrieved from the I/B/E/S database. Out of these dates, we construct a dummy variable *EDAY*, which is equal to one on earnings announcement dates for firm *i*, otherwise zero.

Non-earnings scheduled events are retrieved from the Bloomberg Corporate Events Calendar. Most common non-scheduled events include investor conferences, shareholder meetings, and corporate access (Ben-Rephael et al., 2021). Out of these dates, we construct a dummy variable *NESEDAY*, which is equal to one on non-earnings scheduled events dates for firm *i*, otherwise zero.

Additionally, in line with Ben-Rephael et al. (2021), we construct value-weighted averages of *EDAY* and *NESEDAY* for a given date using all firms in our sample, into the variables *AGG_EDAY* and *AGG_NESEDAY*, respectively.

Macroeconomic announcement dates are retrieved from Bloomberg. Following Ben-Rephael et al. (2021), we consider the five macro announcement events that garner most attention from Bloomberg users. Bloomberg assigns a *relevance score* based on the number of alerts for an announcement added by users. The five macro announcement events with highest relevance scores are: (1) Federal Market Committee rate decisions (FOMC); (2) nonfarm payrolls (NFP); (3) the "advance" forecast of the U.S. Gross Domestic Product (GDP); (4) the Producer Price Index (PPI); and (5) the Institute for Supply Management Manufacturing Index (ISM). For each of these announcement events, we construct a dummy variable, which is equal to one on announcement dates, and zero otherwise. We then construct a summary dummy variable *MACRO*, which is equal to one if any of the five dummy variables for each announcement event is equal to one, otherwise zero.

**IV. Stock Price Non-Synchronicity**

In this paper, the main dependent variable is *stock price non-synchronicity* (NS). Following the intuition in Morck, Yeung, and Yu (2000), the construction of NS is a logistic transformation of $R^2$. The transformation ensures that we obtain a continuous variable that roughly follows a normal distribution, see Appendix B for a visual representation of the measure. The formal definition of *NS* for security *i* is given as:

$$NS_{i,t} = \ln\left[\frac{(1-R_{i,t}^2)}{R_{i,t}^2}\right] \qquad \text{(Eq. 2)}$$

We transform the inverse coefficint of determination $R^2$ for a given firm, by the natural logarithm. To understand how this measure captures non-synchronicity, consider the relationship between $R_i^2$ and the idiosyncratic risk, as given by firm-specific volatility. Ri2depends on both the standard deviation of the error term of stock *i*, $\sigma_{\varepsilon,i}$, as well as the total volatility of stock returns, $\sigma_{r,i}$:

$$R_{i,t}^2 = 1 - \frac{\sigma_{\varepsilon,i}}{\sigma_{r,i}} \qquad \text{(Eq. 3)}$$

Modifying the relationship in *Equation 2* then gives the relationship between idiosyncratic risk and price non-synchronicity:

$$NS_{i,t} = \ln\left[\frac{\sigma_{\varepsilon,i}}{\sigma_{r,i} - \sigma_{\varepsilon,i}}\right] \qquad \text{(Eq. 4)}$$

As *Equations 2-4* show, *NS* can be regarded as a modification of absolute idiosyncratic volatility. In other words, an increase in idiosyncratic volatility also entails an increase in the non-synchronicity of security *i*.

In correspondence to prior studies (Dasgupta et al., 2010; Ferreira, Ferreira, and Raposo, 2011; Kan and Gong, 2018), this paper uses the Fama-French three factor model to estimate $R^2$. We obtain the following model from Fama and French (1993):

$$R_{it} = R_f + \beta_{MKT,i} MKT_T + \beta_{SMB,i} SMB_T + \beta_{HML,i} HML_T + \varepsilon_{it} \qquad \text{(Eq. 5)}$$

In this model, the following notations and variables are included: the firm-specific return ($R_i$), the risk-free rate ($R_f$), the market return premium ($MKT_T$), the small-minus-big company factor ($SMB_T$), and the high-minus-low book-to-market ratio factor ($HML_T$). For each of these variables, there is a corresponding beta $\beta$, which measures the exposure for each of the factors.

As per the framework provided by Fama and French (1993), $MKT_T$ factors systemic exposure to the market, $SMB_T$ factors the performance of small versus big firms, and $HML_T$ factors the performance of high versus low book-to-market firms.

## 2.3 Summary Statistics

The basic properties of the main variables used in this paper are reported in Table I. It shows that an average firm in our sample experiences a factor increase of 1.178 in *non-synchronicity*. Backtracking the mean $R^2$, this means that *non-synchronicity* explains roughly 23.54% of an average firm's returns.

The average firm in our sample experiences an information consumption shock, *AIA*, from institutional investors on 6.9% of all trading days. For retail investors, the average information consumption shock, *ARA*, corresponds to 4.2% of all trading days.

For information events variables, the average stock in our sample has an earnings announcement event, *EDAY*, on 1.8% of all trading days, or roughly 4.5 times per year. Non-scheduled event days, *NESEDAY*, are more frequent, 2.6% of all trading days or roughly 6.5 times per year. The mean for the value-weighted average of earnings announcement days, *AGG_EDAY*, in our sample, is 2%, meaning that 2% of the combined market value in our sample has earnings announcements on a given day. Similarly, the mean of the value-weighted average of non-earnings scheduled events days, *AGG_NESEDAY*, is 5.9%, meaning that, on a given day, 5.9% of the combined market value has a non-earnings scheduled event. Further, we see an average of macroeconomic announcement days, *MACRO*, on 16.7% of all trading days, or roughly 42 times per year.

The average (median) firm in our sample has a size of \$1.47 (\$1.3) billion. On average, \$57.8 million worth of shares is traded per day, for a given stock. The mean (median) return for a given stock in our sample is 12.18 (7.72) bps. Finally, the average firm in our sample has a leverage ratio of 0.205.

***Table I. Summary Statistics***

This table reports summary statistics for our main independent variable, abnormal institutional attention (AIA), and our main dependent variable, non-synchronicity (NS). Other selected variables are also reported. All variables are from the period, February 2010 to December 2017. The full sample includes stocks that appeared in the Russell 3000 index over our sample period, we also require the stock to have nonmissing AIA, book-to-market ratio, and a price of at least \$5. These filters combined with the original sample gives us a sample of 3,624,621 observations across 2464 unique stocks. For variable definitions, see Appendix A. Firms refer to the number of unique firms in our sample. The mean, median and SD refer to the cross-sectional average, median, and standard deviation of the firms' time-series averages.

| *Variable* | *Mean* | *Median* | *SD* | *N* | *n* |
|---|---|---|---|---|---|
| *Dependent variables* | | | | | |
| *NS* | 1.178 | 0.968 | 1.104 | 2464 | 3,624,621 |
| *ILLIQ* | 2.876 | 0.191 | 15.316 | 2464 | 3,624,621 |
| *AIA* | 0.069 | 0.045 | 0.077 | 2464 | 3,624,621 |
| *Independent variables* | | | | | |
| *ARA* | 0.042 | 0.042 | 0.014 | 2464 | 1,612,689 |
| *EDAY* | 0.018 | 0.019 | 0.005 | 2464 | 68,143 |
| *NESEDAY* | 0.026 | 0.022 | 0.023 | 2464 | 97,981 |
| *AGG_EDAY* | 0.020 | 0.006 | 0.004 | 2464 | 3,624,621 |
| *AGG_NESEDAY* | 0.059 | 0.040 | 0.007 | 2464 | 3,624,621 |
| *MACRO* | 0.167 | 0.176 | 0.019 | 2464 | 625,561 |
| *Leverage* | 0.205 | 0.148 | 0.213 | 2464 | 78,156 |
| *lnSize* | 21.110 | 20.992 | 1.666 | 2464 | 3,624,621 |
| *BM* | 0.669 | 0.468 | 9.743 | 2464 | 78,156 |
| *Ret (bps)* | 12.182 | 7.721 | 46.217 | 2464 | 3,624,621 |
| *DVOL ($mn)* | 57.774 | 11.960 | 191.985 | 2464 | 3,624,621 |

# 3. Research Questions and Methodology

In the section below, we aim to develop the hypotheses and the methodology for answering the research questions, with a basis on the data sources and sample construction provided in the previous section.

## 3.1 Research Questions

As the main investigation of this paper - the relationship between investor information consumption and stock price informativeness - is based on the paper of Ben-Rephael et al. (2021), this paper must first test whether AIA is a suitable proxy for investor consumption. For this purpose, this paper includes the following research question:

**Hypothesis I:** *Individual firms' earnings announcements and non-earnings scheduled events, related firms' earnings announcements and non-earnings scheduled events, and macro announcements, are determinants of institutional attention.*

This hypothesis states that the components of the AIA variable are significant, and, thus, strong enough to be determinants of institutional attention. More specifically, these components are: *individual firms' earnings announcements*, *related firms' earnings announcements*, *and macro announcements*. Since this paper aims to strengthen the novel literature of investor information consumption, this hypothesis is a natural research question to confirm whether Ben-Rephael et al. (2021)'s findings are replicable.

**Hypothesis II:** *Investors' information consumption of an individual stock will positively affect the price non-synchronicity of that stock, for institutional investors.*

Our second hypothesis states that there is a positive correlation between the institutional information consumption of a stock, and its *price non-synchronicity*, or, in other words, a negative correlation between the institutional information consumption of a stock and its *price synchronicity*. This hypothesis is concerned with the relationship between investor information consumption and stock price informativeness. Using price non-synchronicity as a proxy for price informativeness, and AIA as a proxy for institutional investor information consumption, this hypothesis tests whether there is a correlation between AIA and price non-synchronicity, and, if applicable, the sign of the correlation.

**Hypothesis III:** *Investors' information consumption of an individual stock will positively affect the price non-synchronicity of that stock, for retail investors.*

The third hypothesis of this paper is similar to its second, with the exchange of institutional investors for retail investors. It states that there is a positive correlation between the retail information consumption of a stock, and its *price non-synchronicity*, or, in other words, a negative correlation between the retail information consumption of a stock and *price synchronicity*. Using price non-synchronicity as a proxy for price informativeness, and ARA (an account of this variable construction is provided above) as a proxy for retail investor consumption, this hypothesis tests whether there is a correlation between ARA and price non-synchronicity, and, if applicable, the sign of the correlation.

### 3.2 Methodology

To investigate and test the research questions provided above, this paper uses a variety of methods to produce robust results and interpretations, with a basis on the paper of Ben-Rephael et al. (2021). This section aims to describe the methodology used to test the hypotheses.

This natural first step of this paper is to follow the methodology provided by Ben-Rephael et al. (2021) to compute for AIA, whilst answering Hypothesis I. To examine the drivers of institutional information consumption, we conduct Logit panel regressions, where AIA is regressed on variables constructed of news releases on firm, market, and macroeconomic levels. Similar to Ben-Rephael et al. (2021), we include day-of-week dummy effects, control variables such as firm characteristics (absolute returns, size, book-to-market, firm beta, leverage ratio, etc.), and double clustered standard errors at the firm level and date level. Day-of-week fixed effects is used to capture eventual seasonality in attention (DellaVigna and Pollet, 2009; Liu and Peng, 2015). Double clustered standard errors negate the issue of correlated residuals over time - and allows for consideration of only the significant variables.

However, this paper lacks the news coverage provided by RavenPack, which is used in Ben-Rephael et al. (2021), due to a lack of access to this database. This entails a minor weakness in the regression coefficients, and, subsequently, the results and interpretations thereof, as a consequence of potential omitted variables.

For Hypothesis II and Hypothesis III, this paper uses the Fama-French three factor model to estimate $R^2$ to compute price non-synchronicity, as per the methodology used in prior studies (Dasgupta et al., 2010; Ferreira, Ferreira, and Raposo, 2011; Kan and Gong, 2018). Further, we use the methodology developed by Morck et al. (2000), to define our dependent variable *NS* for price non-synchronicity, as a logistic transformation of $R^2$.

Further, to test for Hypothesis II and Hypothesis III, this paper uses similar regressions specifications as Ben-Rephael et al. (2021). These similarities include the same usage of AIA (for Hypothesis II), information events, and control variables, with the exception of the data extracted from RavenPack. However, following the methodology in Kan and Gong (2018), this paper uses *NS* as a dependent variable, instead of stock price returns as used in Ben-Rephael et al. (2021).

To answer Hypothesis III, this paper uses Daily Google Search Volume Index data (DSVI). Following Da, Engelberg, and Gao (2011), a logarithmic transformation is made by taking the natural log of the ratio of DSVI, and the average DSVI from the past month. This allows for capturing abnormal retail attention. In accordance with Ben-Rephael et al. (2021), ARA is then constructed using this data, and in the same way as AIA is constructed. This ensures comparison capabilities between ARA and AIA. To answer Hypothesis III, this paper regresses for non-synchronicity, with ARA as the main independent variable.

Further, to account for eventual "lags" of significant effects on price informativeness, due to information inertia (Illeditsch, Ganguli, and Condie, 2021), this paper conducts secondary panel regressions with an artificial "lag" to the AIA variable. This takes into account information inertia as well as the possibility of a "lag" between information consumption and price reaction.

To produce robust results and interpretations, this paper includes robustness checks. In accordance with Li, Rajgopal, and Venkatachalam (2014), this paper introduces Amihud (2002)'s illiquidity ratio as an alternative measure of price informativeness. In Amihud's illiquidity ratio, we construct time intervals of two trading weeks, or ten days. This balance ensures that not too many observations per firm are lost, while some stability is still inferred. The inclusion of this analysis and comparison with price non-synchronicity is warranted through a wide usage in prior literature (Kelly, 2014; Li et al., 2014; Kan and Gong, 2018; Gassen, Skaife, Veenman, 2020).

Further, this paper includes a placebo test, to examine whether the results are robust or driven by noise. The placebo test is constructed as follows: a random dummy variable is created, which follows a binomial distribution and is equal to one with a similar probability to that of AIA being equal to one. This test checks if the relation of AIA to non-synchronicity is of mere chance and in favor of the null hypothesis - no connection between AIA and non-synchronicity. A further development on this test is by creating a completely random variable with the same distribution as news readership data in Bloomberg. Unfortunately, that data is out of the scope of access of this paper, and we rely on this rudimentary randomization for this particular analysis.

To address multicollinearity concerns, this paper includes a multicollinearity analysis. A correlation matrix is constructed, containing the main variables used in this paper. This matrix discloses if there are strong correlations that might imply statistical errors or paradoxes (such as Simpson's paradox). Additionally, this paper includes a tolerance test and variance inflation factor test, *VIF*. These tests provide indicators of whether multicollinearity is an issue in our specifications and results.

# 4. Results and Analysis

In this section, we present the results of this paper, as well as interpretations of the results. We offer discussion on the implications on our hypotheses, and give an indication on how robust our results and interpretations are.

The section is structured as per the manner in the methodology section. First, an account of the *Determinant of Institutional Attention* is given. This is followed by *Non-Synchronicity and Information Consumption*, the latter being subdivided into terms of institutional investors and terms of retail investors.

## 4.1 Determinants of Institutional Attention

The effort to answer Hypothesis I, proceeds by determining the drivers of institutional attention. For a complete breakdown of the construction of the variables used in the regressions specifications, see *Sample Construction and Data Sources*.

Our main analysis follows the following logistic panel regression model:

$$AIA_{it} = \beta_1 EDAY_{it} + \beta_2 NESEDAY_{it} + \beta_3 AGG\_EDAY_{it} + \beta_3 AGG\_NESEDAY_{it} + \beta_5 MACRO_{it} + \beta_X X_{it} + \alpha_t + \varepsilon_{it} \tag{Eq. 6}$$

The model uses three proxies for information events, following Ben-Rephael et al. (2021). The variable $X$ is a vector of control variables. Day-of-week fixed effects are used to capture eventual seasonality in attention.

The results from the Logit panel regressions are presented in *Table II*. We can observe a high explanatory value of the firm-specific information events, $EDAY$ and $NESEDAY$. Of these two, $EDAY$ has the highest coefficient for all models. We cannot observe a significant effect of *MACRO* as a driver for institutional attention. We find a significant effect of $AGG\_EDAY$ and $AGG\_NESEDAY$ as drivers of institutional attention.

Our results infer roughly similar interpretations of the drivers of institutional attention, as in Ben-Rephael et al. (2021). However, one point of differentiation is that we, in contrast to Ben-Rephael et al. (2021), do not find an effect of macro announcements on institutional attention. Even in Ben-Rephael et al. (2021), the effect of macro announcement is only captured at the lowest significance level, $p<0.05$, which suggests that the evidence was low for this effect. Our results indicate that firm-specific information events drive institutional information consumption. Of the two firm-specific information events, earnings announcements and non-earnings scheduled events, the former drives institutional attention to a larger extent. Further, we find that other firms' information events also have a significant effect for the institutional attention of a given firm, both for other firms' earnings announcements and non-earnings scheduled

## *Table II. Determinants of Institutional Attention*

This table reports results from the Logit panel regressions of the institutional attention variable (AIA) measure that we obtained from Bloomberg on measures of scheduled information events, both from firm-specific information events, macroeconomic information events, and other firms' scheduled events. For variable definitions, see Appendix A. Model 1 includes firm-specific scheduled events, through dummies for earnings announcements, EDAY, and non-earnings scheduled events, NESEDAY. Model 2 includes the summary variable for macroeconomic announcements, MACRO, which is equal to one if any of the five major macroeconomics announcements have been made on that day. For model 3, we add the market level value-weighted average of both non-earnings scheduled events, AGG_NESEDAY, and earnings announcements, AGG_EDAY. For model 4, 5, and 7 we also introduce separate dummies for each of the five major macroeconomic announcements: Federal Open Market Committee (FOMC) rate decisions, Nonfarm payrolls (NFP), Producer Price Index (PPI), ISM manufacturing index (ISM), and advance estimate for GDP, (GDP). In the last two models, 6-7, we also include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of the firm's book-to-market ratio (lnBM), the absolute returns for a firm (abs_ret), the firm's CAPM beta (b_mkt), as well as the firm's leverage ratio (Leverage), i.e., the ratio between the firm's long-term debt and its total assets. All specifications include day-of-week fixed effects. The sample includes 3,624,261 observations. Additionally, the standard errors are clustered by both firm and date.

| VARIABLES | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|
| EDAY | 0.3938 *** | 0.3937 *** | 0.3914 *** | 0.3936 *** | 0.3913 *** | 0.3536 *** | 0.3537 *** |
| | (0.0077) | (0.0077) | (0.0075) | (0.0077) | (0.0075) | (0.0071) | (0.0071) |
| NESEDAY | 0.1230 *** | 0.1231 *** | 0.1235 *** | 0.1231 *** | 0.1235 *** | 0.0873 *** | 0.0872 *** |
| | (0.0053) | (0.0053) | (0.0052) | (0.0053) | (0.0052) | (0.0039) | (0.0039) |
| MACRO | | 0.0029 | 0.0027 | | | 0.0009 | |
| | | (0.0018) | (0.0018) | | | (0.0021) | |
| AGG_EDAY | | | 0.0765 ** | | 0.0755 ** | 0.1628 *** | 0.1643 *** |
| | | | (0.0287) | | (0.0288) | (0.0280) | (0.0280) |
| AGG_NESEDAY | | | 0.0006 | | -0.0002 | 0.0543 *** | 0.0525 *** |
| | | | (0.0155) | | (0.0155) | (0.0157) | (0.0158) |
| FOMC | | | | 0.0073 | 0.0062 | | 0.0014 |
| | | | | (0.0040) | (0.0040) | | (0.0048) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PPI | | | | 0.0010<br>(0.0035) | 0.0011<br>(0.0034) | | 0.0026<br>(0.0038) |
| NFP | | | | 0.0037<br>(0.0043) | 0.0042<br>(0.0043) | | 0.0044<br>(0.0047) |
| ISM | | | | 0.0031<br>(0.0035) | 0.0030<br>(0.0035) | | -0.0022<br>(0.0040) |
| GDP | | | | -0.0012<br>(0.0031) | -0.0015<br>(0.0031) | | -0.0028<br>(0.0037) |
| *Adjusted $R^2$* | 0.0480 | 0.0480 | 0.0484 | 0.0480 | 0.0484 | 0.1124 | 0.1124 |
| *N* | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 |
| *n* | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 |
| *Day-of-week FE?* | YES | YES | YES | YES | YES | YES | YES |
| *Other controls?* | | | | | | YES | YES |

Standard errors are double-clustered at the firm and date level. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.0$

events. This suggests information spillover effects on specific firms. It is in line with intuition, considering that firm announcements can reveal clues about entire industries. There is academic support (Berman and Pfleeger, 1997; Fort, Haltiwanger, Jarmin, and Miranda, 2013) for our results in that some firms are sensitive to cyclicality of the economy, while others are secular from the wider economy, which explains some of the variation captured in *Table II*.

This analysis provides evidence in part in favor of Hypothesis I, by confirming that individual firms' earnings announcements and non-earnings scheduled events, as well as related firms' earnings announcements and non-earnings scheduled events are determinants of institutional attention. However, our findings do not support that macro announcements are determinants of institutional attention, which is in contrast to Hypothesis I. This analysis also tells us that we have similar characteristics as the dataset in Ben-Rephael et al. (2021), albeit the difference in significance of the macroeconomic announcement variable.

The limitations of our results should be noted, and are mainly attributed to our lack of access to RavenPack news coverage data. Compared to Ben-Rephael et al. (2021), we lack a variable for newsdays. They base this variable on news from the Dow Jones Newswire, and find a significant positive effect on institutional attention. Failing to include this variable, therefore implies omitted variable bias into our model-estimation. Even though a full analysis for this bias should be based on on the correlation between newsdays and other firm-specific information events (earnings announcements or non-earnings scheduled events), the need of this analysis is drastically reduced since Ben-Rephael et al. (2021) constructed the news variables as dummy variables, and excluded any news released on earnings announcement days or non-earnings scheduled event days. By doing this, the correlation between the news variables and information event variables is heavily reduced. Still, the estimated coefficients in our regressions are higher than those of Ben-Rephael et al. (2021), which is likely attributed to the omitted variable bias.

## 4.2 Non-Synchronicity and Information Consumption

### *A. Non-Synchronicity and Institutional Information Consumption*

The effort to answer Hypothesis II, proceeds by investigating the relation between institutional investor information consumption and stock price non-synchronicity. A complete breakdown of the construction of the variables used in the regression specifications are given in *Sample Construction and Data Sources*.

Our main analysis follows the following panel regression model:

$$NS_{it} = \beta_1 AIA_{it} + \beta_2 EDAY_{it} + \beta_3 AIA * EDAY + \beta_4 NESEDAY_{it} + \beta_X X_{it} + \alpha_t + \varepsilon_{it} \quad \text{(Eq. 7)}$$

The model uses three proxies for information events, following Ben-Rephael et al. (2021). The variable *X* is a vector of control variables. Date fixed effects are used to account for eventual development of information consumption over time. In the last model specification *Model 8*, we also include sector fixed effects.

The results from the panel regressions are reported in *Table III*. While regressing for non-synchronicity, we observe a high explanatory value of *AIA*. A spike in AIA increases non-synchronicity by more than 0.23 with significance levels robust over all models. Further, we find a significant, negative correlation of earnings announcement days on price non-synchronicity. The observations of the interaction term *AIAxEDAY* are not economically nor statistically significant for any of the regression models. We also observe a statistically significant positive correlation between non-synchronicity and *NESEDAY*.

Our results infer several interpretations. First, *AIA* does economically and statistically significantly affect non-synchronicity, meaning spikes in institutional investor consumption of a firm, increases its price non-synchronicity, and, thus, price informativeness. Although not completely comparable, this result is in line with Ben-Rephael et al. (2021), which finds a positive correlation between *AIA* and a firm's returns. The intuition is that an immediate increase in returns of a firm might entail an increase of firm-specific (private) information in its price composition. This holds under the assumption that a market model cannot explain the firm-specific return spikes in conjunction with *AIA* spikes.

Second, we take notice of the negative correlation of earnings announcement days on non-synchronicity, and we compare these observations with those of the interaction term. We infer from the results of *Model 8* that the net effect on non-synchronicity, on days when a given firm *i* experiences both abnormal institutional attention and earnings announcement, is then 0.125, all other things equal. For the other regressions in the model, this effect is even higher. This suggests that when other information events than earnings announcements trigger abnormal institutional attention, this entails a stronger effect on non-synchronicity. This finding has some support from both Roll (1988) and Dasgupta et al. (2010), who both suggest that earnings announcement events have no significant effect on price informativeness.

Third, the statistically significant effect of non-earnings scheduled events imply a driver for non-synchronicity. However, in comparison with abnormal institutional attention, the effect is of drastically less economic significance, rendering the implication of low significance. The effect is shaped as a positive correlation. This follows intuition, as the most common categories of non-earnings scheduled events are investor conferences, shareholder meetings, and corporate access, all of which may reveal significant information to the participants and market. This view has support from Brickley (1986), who showed that abnormal returns can arise around proxy statements due to the potential information such events produce. Intuitively, one might speculate that this type of informative events might educate investors to consume information in conjunction with non-earnings scheduled events.

Moreover, the specification *Model 8* in our regressions reveals the effect of sector fixed effects. The results suggest that different sectors experience different effects in terms of non-synchronicity, and that some sectors are more prone to abnormal institutional information consumption, and some less. This interpretation is supported by both Bhushan (1989) and Gompers and Metrick (2001).

This analysis provides evidence in favor of Hypothesis II, by confirming that institutional investors' information consumption of a given stock positively affects the price non-synchronicity, and, thus, price informativeness of that particular stock. There are both economical and statistical significance to the rejection of the null hypothesis, which states that there is no correlation between price institutional investor information consumption and price non-synchronicity. Further, we learn that earnings announcement events negate the effect of institutional investor information consumption on price non-synchronicity.

## *Table III. Abnormal Institutional Attention and Non-Synchronicity*

This table reports results from the panel regression of non-synchronicity (NS) on abnormal institutional attention (AIA), as well as other informational events, both firm-specific and non-firm-specific. For variable definitions, see Appendix A. Model 1 uses only the institutional investor attention measure (AIA). Models 2 and 3 introduce firm earnings announcements (EDAY) and non-earnings scheduled events (NESEDAY), respectively. Models 4 and 5 both include the market wide value-weighted averages of earnings announcements (AGG_EDAY) and, similarly, for the non-earnings scheduled events (AGG_NESEDAY). To facilitate comparison, the value-weighted variables are multiplied by 100. All specifications include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of firm book-to-market ratio (lnBM), the firm leverage ratio (Leverage), i.e., the ratio of long-term debt to total assets, the percentage of shares outstanding held by institutions (inshtold), and the market beta from the CAPM (b_mkt). All specifications include date fixed effects and specification 7 also includes sector fixed effects, there are 11 sectors. The sample includes 3,624,261 observations. Additionally, all standard errors are clustered by both firm and date.

| VARIABLES | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| AIA | 0.2724 *** | 0.2845 *** | 0.2850 *** | 0.2810 *** | 0.2800 *** | 0.2801 *** | 0.2806 *** | 0.2345 *** |
| | (0.0111) | (0.0115) | (0.0118) | (0.0117) | (0.0114) | (0.0114) | (0.0116) | (0.0105) |
| EDAY | | -0.1144 *** | -0.1107 *** | -0.1107 *** | -0.1289 *** | -0.1109 *** | -0.1196 *** | -0.1004 *** |
| | | (0.0076) | (0.0098) | (0.0098) | (0.0089) | (0.0098) | (0.0100) | (0.0077) |
| AIA x EDAY | | | -0.0085 | -0.0043 | -0.0221 | -0.0033 | -0.0204 | -0.0091 |
| | | | (0.0121) | (0.0121) | (0.0135) | (0.0120) | (0.0132) | (0.0122) |
| NESEDAY | | | | 0.0945 *** | 0.0942 *** | 0.0744 *** | 0.0747 *** | 0.0502 *** |
| | | | | (0.0092) | (0.0092) | (0.0102) | (0.0101) | (0.0092) |
| AGG_EDAY | | | | | -0.2882 *** | | -0.2757 *** | -0.1574 ** |
| | | | | | (0.0583) | | (0.0544) | (0.0555) |
| AGG_NESEDAY | | | | | | -0.1682 *** | -0.1628 *** | -0.0752 |
| | | | | | | (0.0494) | (0.0482) | (0.0499) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Adjusted $R^2$* | 0.4798 | 0.4800 | 0.4800 | 0.4802 | 0.4802 | 0.4803 | 0.4803 | 0.5367 |
| *N (Sectors)* | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 (11) |
| *n* | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 |
| *Day FE? (Sector FE?)* | YES | YES | YES | YES | YES | YES | YES | YES (YES) |
| *Other controls?* | YES | YES | YES | YES | YES | YES | YES | YES |

Standard errors are double-clustered by firm and date. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Even though this analysis provides evidence in favour of our hypothesis, it is important to remember that the interpretation of the results is limited by potential limitations. It is mainly the potential of omitted variables and us not using firm fixed effects that is cause for concern. (We wanted to include firm fixed effects, but the size of the dataset made it too computationally intense.) It may therefore be that there are some constant firm characteristics that we are not taking into account into our analysis that biases our results. Still, the tests we conduct below tries to remedy this problem somewhat.

### *B. Time-lagged Abnormal Institutional Attention (AIA)*

In a continued effort to answer Hypothesis II, this paper proceeds by investigating the relation between institutional investor information consumption and stock price non-synchronicity, in a setting with a post-earnings announcement drift.

As post-earnings announcement drifts have implied effects on price informativeness, due to an inertia following news announcements, we introduce artificial "lags" to our panel regressions, to test for this factor.

Our analysis follows the following panel regression model:

$$\begin{aligned} NS_{it} = \beta_1 AIA_{i(t-(1:5))} + \beta_2 EDAY_{i(t-(1:5))} + \beta_4 NESEDAY_{i(t-(1:5))} \\ + \beta_5 AGG\_EDAY_{it} + \beta_X X_{it} + \alpha_t + \varepsilon_{it} \end{aligned} \qquad \text{(Eq. 8)}$$

The model uses three proxies for information events, following Ben-Rephael et al. (2021). The variable *X* is a vector of control variables. Date fixed effects are used to account for eventual development of information consumption over time. This model differs from the model underlying *Table III*, by including "lags" of the *AIA* variable of one up to five days.

The results from the panel regressions with the "lagged" *AIA* are presented in *Table IV*. Similar to *Table III*, we observe a high explanatory value of *AIA* for all models in our regressions. We can also observe a decrease of the adjusted $R^2$ for each additional lag.

These results imply that the effect of *AIA*, and, thus, a shock in institutional attention, is persistent over a trading week. Simultaneously, our results imply that changes in price non-synchronicity, and, thus, price informativeness are persistent over time. We show, then, that Illeditsch, Ganguli, and Condie (2021)'s suggestions that risk and ambiguity lead to inefficiency in information processing, and thus, an inertia or "lag" are two-fold. Our results do not indicate any delays in the effect of *AIA* on non-synchronicity, suggesting information inefficiency or inertia does not influence the delay of the effect significantly. However, we can observe that the effect of *AIA* is persistent over a longer period of time, suggesting that information inefficiency or inertia are influencing factors in slowing the non-synchronicity in resetting from the effect of the information consumption shock. This finding is in accordance with related literature on post-earnings announcement drifts, which show that abnormal returns can follow earnings announcements for an extended period of time.

The explanations for this effect are debated. Liang (2003) attributes this effect to market participant bias. Following the explanation offered by Liang (2003), we

### *Table IV. Time-lagged Abnormal Institutional Attention and Non-Synchronicity*

This table reports results from the panel regression of non-synchronicity (NS) on lagged abnormal institutional attention (AIA), as well as other lagged informational events. For variable definitions, see Appendix A. Model 1 is the same model as Model 7 in Table III, it is intended to work as a comparison. From Model 2 and onward to Model 6, the attention variable (AIA) and the other firm-specific information event variables (EDAY and NESEDAY) are lagged by one additional day for each Model. This means that Model 6 reports the results from the panel regression of non-synchronicity (NS) on a five-day lagged consortium of information events and attention. To facilitate comparison, the value-weighted average of earnings announcing firms, AGG_EDAY, har been multiplied by 100, as we did in Table III. The same is done for the value-weighted average of firms with non-earnings scheduled events, AGG_NESEDAY. All specifications include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of firm book-to-market ratio (lnBM), the firm leverage ratio (Leverage), i.e., the ratio of long-term debt to total assets, the percentage of shares outstanding held by institutions (insthold), and the market beta from the CAPM (b_mkt). All specifications include date fixed effects. The sample includes 3,624,261 observations. Additionally, the standard errors are clustered by firm and date.

| Variables | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| AIA | 0.2792 *** | 0.2702 *** | 0.2698 *** | 0.2703 *** | 0.2705 *** | 0.2711 *** |
| | (0.0113) | (0.0113) | (0.0113) | (0.0113) | (0.0113) | (0.0113) |
| EDAY | -0.1279 *** | -0.1213 *** | -0.1196 *** | -0.1184 *** | -0.1156 *** | -0.1149 *** |
| | (0.0087) | (0.0075) | (0.0075) | (0.0075) | (0.0075) | (0.0075) |
| NESEDAY | 0.0748 *** | 0.0893 *** | 0.0904 *** | 0.0903 *** | 0.0898 *** | 0.0889 *** |
| | (0.0101) | (0.0090) | (0.0091) | (0.0091) | (0.0091) | (0.0091) |
| AGG_EDAY | -0.2673 *** | -0.1707 *** | -0.2672 *** | -0.2958 *** | -0.2962 *** | -0.2978 *** |
| | (0.0516) | (0.0422) | (0.0504) | (0.0535) | (0.0542) | (0.0539) |
| AGG_NESEDAY | -0.1631 *** | -0.2695 *** | -0.2812 *** | -0.2805 *** | -0.2778 *** | -0.2776 *** |
| | (0.0483) | (0.0495) | (0.0504) | (0.0503) | (0.0501) | (0.0502) |
| *Adjusted* $R^2$ | 0.4803 | 0.4801 | 0.4801 | 0.4801 | 0.4799 | 0.4802 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *N* | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 |
| *n* | 3,624,621 | 3,624,620 | 3,624,619 | 3,624,618 | 3,624,617 | 3,624,616 |
| *Date FE?* | YES | YES | YES | YES | YES | YES |
| *Other controls?* | YES | YES | YES | YES | YES | YES |

Standard errors are double clustered by firm and date. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

can infer that this effect on non-synchronicity is attributed to investors' overconfidence in their own private information, or a general asymmetry between confidence and reliable information within the market. Similarly, Bernard and Thomas (1989) find evidence that the price-earnings announcement drift is a delayed price response to information, and an inefficiency in the inference of future earnings from current earnings announcements. Both of these papers argue for hypotheses that are reconcilable with our results. However, we remain humble that our results are merely indications, and more robust tests and experiments have to be conducted to reach a conclusion on the nature of the persistence of the information consumption effect on price non-synchronicity.

### *C. Retail Attention as an Explanation for Non-Synchronicity*

The effort to answer Hypothesis III, proceeds by investigating the relation between retail investor information consumption and stock price non-synchronicity. A complete breakdown of the construction of the variables used in the regression specifications are given in *Sample Construction and Data Sources*.

Our main analysis is based on the following panel regression model:

$$NS_{it} = \beta_1 ARA_{it} + \beta_2 EDAY_{it} + \beta_3 ARA * EDAY + \beta_4 NESEDAY_{it} + \beta_X X_{it} + \alpha_t + \varepsilon_{it} \quad \text{(Eq. 9)}$$

The model uses three proxies for information events, following Ben-Rephael et al. (2021). The variable *X* is a vector of control variables. Date fixed effects are used to account for eventual development of information consumption over time. The panel regressions run are similar to those underlying *Table III*, however, with exchanging *AIA* for *ARA* - abnormal retail attention. In the last model specification *Model 8*, we also include sector fixed effects. The results of the panel regressions are presented in *Table V*. We can observe a none or statistical significance at the lowest level, $p<0.05$, for *ARA* after including other information event variables. For *Models 3-8*, *ARA* does not surpass 0.008.

We can interpret from the results that abnormal retail attention has no economic, and no or low statistical significance as an explanatory variable for non-synchronicity. Therefore, we find no support for Hypothesis III, in that retail investor information consumption of a stock will positively affect the non-synchronicity of that particular stock. We can therefore not reject the null hypothesis, that there is no effect on price non-synchronicity by retail investor attention. This is in line with Ben-Rephael et al. (2021), who find no systematic implications whatsoever of retail information consumption. We can, however, draw weak inferences about the relation between retail information consumption and price non-synchronicity - retail investors are unlikely to increase price informativeness, compared to institutional investors. This, too, is in line with previous literature (Ben-Rephael, Da, and Israelsen, 2017).

It should be noted that our results are limited by data coverage, which is less than that of Ben-Rephael et al. (2021). This entails less robust results and interpretations, thereof. It does, however, indicate similar interpretations as prior studies, and is not in contrast to them.

## *Table V. Retail attention and Non-Synchronicity*

This table reports results from the panel regression of non-synchronicity (NS) on abnormal retail attention (ARA), as well as other informational events, both firm-specific and non-firm-specific. For variable definitions, see Appendix A. Model 1 uses only the retail investor attention measure (ARA). Models 2 and 3 introduce firm earnings announcements (EDAY) and non-earnings scheduled events (NESEDAY), respectively. Models 4 and 5 both include the market wide value-weighted averages of earnings announcements (AGG_EDAY) and, similarly, for the non-earnings scheduled events (AGG_NESEDAY). To facilitate comparison, the value-weighted variables are multiplied by 100. All specifications include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of firm book-to-market ratio (lnBM), the firm leverage ratio (Leverage), i.e., the ratio of long-term debt to total assets, the percentage of shares outstanding held by institutions (insthold), and the firm's CAPM market beta (b_mkt). All specifications include date fixed effects and Model 8 also includes sector fixed effects. Additionally, all standard errors are double clustered by firm and date. Due to data coverage, the sample only includes 1,612,689 daily observations.

| VARIABLES | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| ARA | 0.0156 ***<br>(0.0032) | 0.0155 ***<br>(0.0032) | 0.0080 *<br>(0.0032) | 0.0078 *<br>(0.0032) | 0.0078 **<br>(0.0032) | 0.0077 *<br>(0.0032) | 0.0077 *<br>(0.0032) | 0.0049<br>(0.0029) |
| EDAY | | 0.0050<br>(0.0082) | -0.0146<br>(0.0088) | -0.0142<br>(0.0088) | -0.0290 **<br>(0.0098) | -0.0142<br>(0.0088) | -0.0283 **<br>(0.0097) | -0.0226 ***<br>(0.0066) |
| AIAxEDAY | | | 0.2334 ***<br>(0.0235) | 0.2332 ***<br>(0.0235) | 0.2003 ***<br>(0.0237) | 0.2317 ***<br>(0.0234) | 0.2003 ***<br>(0.0238) | 0.1671 ***<br>(0.0208) |
| NESEDAY | | | | 0.1056 ***<br>(0.0109) | 0.1053 ***<br>(0.0109) | 0.0858 ***<br>(0.0121) | 0.0862 ***<br>(0.0121) | 0.0533 ***<br>(0.0102) |
| AGG_EDAY | | | | | -0.2888 ***<br>(0.0511) | | -0.2757 ***<br>(0.0478) | -0.1519 ***<br>(0.0444) |
| AGG_NESEDAY | | | | | | -0.1858 **<br>(0.0664) | -0.1800 **<br>(0.0655) | -0.1324 *<br>(0.0538) |
| *Adjusted* $R^2$ | 0.4087 | 0.4087 | 0.4088 | 0.4088 | 0.4091 | 0.4091 | 0.4092 | 0.5006 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *N* | | | | | | | | |
| *n* | 1,612,689 | 1,612,689 | 1,612,689 | 1,612,689 | 1,612,689 | 1,612,689 | 1,612,689 | 1,612,689 |
| *Date FE? (Sector FE?)* | YES | YES | YES | YES | YES | YES | YES | YES (YES) |
| *Other controls?* | YES | YES | YES | YES | YES | YES | YES | YES |

Standard errors are double clustered by firm and date. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.0$

# 5. Robustness

In this paper, several robustness checks are conducted in order to produce robust results and interpretations. In the section below, we give account for some of the tests and checks conducted in this paper. It describes the construction of the robustness checks, the results, and the interpretations on our results and conclusions.

## 5.1 Alternative Proxy of Price Informativeness: Amihud's Illiquidity Ratio

As an alternative measure of prince informativeness, this section presents Amihud (2002)'s illiquidity measure. In this section, the construction and results of this test are provided, as well as discussion of the implied results.

Amihud's illiquidity ratio follows Amihud (2002), and is given by the following model:

$$ILLIQ = \frac{1}{D_{it}} \sum_{i=1}^{D_{it}} \frac{|R_{it}|}{DVOL_{it}} \qquad \text{(Eq. 10)}$$

Where $D_{it}$ is the number of days in the time interval, and for which data are available for firm *i* in time *t*, $DVOL_{it}$ is the trading volume in dollars for firm *i* in time *t*, and $RET_{it}$ is the return for firm *i* in time *t*. In our regressions, we use intervals of *two* trading weeks, or *ten* days.

The results are presented in *Table VI*. Regressing for *ILLIQ*, we observe a statistically significant increase of 1.2017 of AIA in *Model 8*, and observe similar positive numbers for all models, controlling for different variables as in previous regressions. Our results show a positive correlation between AIA and Amihud's illiquidity ratio.

These regression results act as a robustness check and in support of previous results of a positive correlation between price non-synchronicity and AIA. It is also in line with prior literature by Li, Rajgopal, and Venkatachalam (2014), as they conclude that a higher *liquidity* ratio implies less informative prices, and, thus, higher *illiquidity* ratio implies more informative prices. Since we observe similar effects of AIA on two proxies of price informativeness - illiquidity and non-synchronicity - this analysis renders our interpretations more robust.

## 5.2 Placebo Test: Random Variable

This section describes the placebo test performed in order to examine whether the results are robust, or noise driven. In this section, the construction and results of this test are provided, as well as discussion of the implied results.

The placebo test is constructed as follows: a random dummy variable is created, which follows a binomial distribution and is equal to one with a probability of 0.0765, which corresponds to the average (mean) of AIA, without the time-average dimension for each firm as presented in *Table I: Summary Statistics*. The regressions in *Table III* are then recreated, with this random variable as a substitute to AIA, and presented in *Table VII*. The regressions are otherwise made using the same specifications as in *Table III*.

Looking at *Table VII*, we can observe similar results for all model specifications. The random variable has no significant effect on the non-synchronicity for a given firm. Notably, we can also observe from *Table VII* that earnings announcement days have negligent effects on non-synchronicity for a given firm, when AIA is not accounted for. We can observe that the signs of the information events variables remain unchanged in this test.

Firstly, these results support our previous results of a positive correlation between AIA and price non-synchronicity. This test provides robustness by confirming that non-synchronicity is driven by AIA and not by noise. Therefore, the probability that the observed relationship between AIA and price non-synchronicity is a product of mere chance, is radically diminished. For this reason, we can interpret this check as supporting the rejection of the null hypothesis, based on the previous regressions of non-synchronicity with AIA as the main independent variable.

Also, this test lets us examine closer the relationship between AIA and EDAY. We see a non-significant or negligent effect of EDAY, when the variable of AIA is not taken into account. However, when both variables are included in the regressions, we see a stronger negative tendency of EDAY. This implies that the effect of earnings announcements is dependent on the variables omitted. Similarly, it implies that earnings announcements often do not surprise investors with unexpected information and when they do, investors tend to consume that information and increase the price informativeness of the stock.

We remain humble of the fact that, even though this test supports the proposition of AIA as a driver for price informativeness, Hypothesis II, it does not negate the potential issues of omitted variables. It does neither provide evidence to reject the possibility of reverse causality, and these two factors should still influence the interpretation of the results.

### 5.3 Multicollinearity Checks

To address any multicollinearity concerns, this section describes the multicollinearity checks performed in order to determine whether there remain issues regarding multicollinearity. In this section, the construction and results of these tests are provided, as well as discussion of the implied results.

The first check includes a correlation matrix. The correlation matrix contains the main variables used in the different regressions in this paper. Further, we test for tolerance and for variance inflation factor (*VIF*), for *Table II*. Moreover, we test for eigenvalues and Condition Index for *Table II*. We also test for VIF for *Table III*.

The results of the correlation matrix can be seen in Appendix C. Among the strongest correlations, we find the market beta and non-synchronicity, with a correlation of -0.48. We also find a correlation between the natural logarithm of market size and trading volume in dollars, of 0.44. The results of the tolerance test and VIF for *Table II* can be seen in Appendix D, *Table D.1*. The results of VIF for *Table III* can be seen in Appendix E. The tolerance test indicates tolerance scores of above 0.7934 for all variables in *Table II*. VIF have scores of below 1.2603 and 1.6064, for *Table II* and *Table III*, respectively. The results of the eigenvalues and Condition Index can be seen in Appendix D, *Table D.2*. We can observe scores of above 0.0541 and below 10.1233 for eigenvalues and Condition Index, respectively.

We understand from the various multicollinearity checks that multicollinearity is not a major issue in our regressions, which renders them more reliable and robust.

## *Table VI. Institutional Attention and Illiquidity*

This table reports results from the panel regression of illiquidity measure (ILLIQ) on abnormal institutional attention (AIA), as well as other informational events, both firm-specific and non-firm-specific. For variable definitions, see Appendix I. Model 1 uses only the institutional investor attention measure (AIA). Models 2 and 3 introduce firm earnings announcements (EDAY) and non-earnings scheduled events (NESEDAY), respectively. Models 4 and 5 both include the market wide value-weighted averages of earnings announcements (AGG_EDAY) and, similarly, for the non-earnings scheduled events (AGG_NESEDAY). To facilitate comparison, the value-weighted variables are multiplied by 100. All specifications include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of firm book-to-market ratio (lnBM), the firm leverage ratio (Leverage), i.e., the ratio of long-term debt to total assets, the percentage of shares outstanding held by institutions (insthold), and the firm's market beta from the CAPM (b_mkt). All specifications include date fixed effects and Model 6 includes sector fixed effects as well, there are 11 sectors. By creating the illiquidity measure, we lose some observations, the regressions are thus based on 3,602,471 observations. Additionally, all standard errors are double clustered by firm and date.

| VARIABLES | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| AIA | 1.2277 *** | 1.3214. *** | 1.2930 *** | 1.2789 *** | 1.2930 *** | 1.2606 *** | 1.2692 *** | 1.2017 *** |
| | (0.1416) | (0.1524) | (0.1482) | (0.1467) | (0.300) | (0.1447) | (0.1456) | (0.1407) |
| EDAY | | -0.8916 *** | -1.0693 *** | -1.0695 *** | -1.2103 *** | -1.0722 *** | -1.2051 *** | -1.1702 *** |
| | | (0.1295) | (0.1885) | (0.1885) | (0.1885) | (0.1887) | (0.2003) | (0.1959) |
| AIAxEDAY | | | 0.4083 * | 0.4229 * | 0.1478 | 0.4414 | 0.1812 | 0.2036 |
| | | | (0.1690) | (0.1692) | (0.1692) | (0.1697) | (0.1697) | (0.1591) |
| NESEDAY | | | | 0.3271 *** | 0.3212 *** | -0.0716 | -0.0671 | -0.0650 |
| | | | | (0.0685) | (0.0685) | (0.0698) | (0.0697) | (0.0687) |
| AGG_EDAY | | | | | -4.4447 *** | | -4.1959 *** | -3.9302 *** |
| | | | | | (0.8315) | | (0.8123) | (0.7901) |
| AGG_NESEDAY | | | | | | -3.3231 *** | -3.2394 *** | -2.8040 *** |
| | | | | | | (0.6077) | (0.5958) | (0.5574) |
| *Adjusted* $R^2$ | 0.1069 | 0.1070 | 0.1070 | 0.1070 | 0.1070 | 0.1070 | 0.1071 | 0.1105 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *N* | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 |
| *n* | 3,602,471 | 3,602,471 | 3,602,471 | 3,602,471 | 3,602,471 | 3,602,471 | 3,602,471 | 3,602,471 |
| *Date FE? (Sector FE?)* | YES | YES | YES | YES | YES | YES | YES | YES (YES) |
| *Other controls?* | YES | YES | YES | YES | YES | YES | YES | YES |

Standard errors are double clustered by both firm and date. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

### *Table VII. Random Variable and Non-Synchronicity*

This table reports results from the panel regression of non-synchronicity (NS) on a random variable (random), as well as other informational events, both firm-specific and non-firm-specific. For variable definitions, see Appendix A. Model 1 uses only the random variable measure (random). Models 2 and 3 introduce firm earnings announcements (EDAY) and non-earnings scheduled events (NESEDAY), respectively. Models 4 and 5 both include the market wide value-weighted averages of earnings announcements (AGG_EDAY) and, similarly, for the non-earnings scheduled events (AGG_NESEDAY). To facilitate comparison, the value-weighted variables are multiplied by 100. All specifications include control variables. These are: the natural logarithm of firm market capitalization (lnSize), the natural logarithm of firm book-to-market ratio (lnBM), the firm leverage ratio (Leverage), i.e., the ratio of long-term debt to total assets, the percentage of shares outstanding held by institutions (insthold), and the firm's market beta from the CAPM (b_mkt). All specifications include date fixed effects. Specification 7 also includes sector fixed effects, there are 11 sectors. The sample includes 3,624,261 observations. Additionally, all standard errors are double clustered by firm and date.

| VARIABLES | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|---|
| random | -0.0010<br>(0.0015) | -0.0012<br>(0.0015) | -0.0012<br>(0.0015) | -0.0012<br>(0.0015) | -0.0012<br>(0.0015) | -0.0012<br>(0.0015) | -0.0012<br>(0.0015) | -0.0001<br>(0.0014) |
| EDAY | | -0.0045<br>(0.0057) | -0.0052<br>(0.0057) | -0.0047<br>(0.0058) | -0.0252 ***<br>(0.0097) | -0.0048<br>(0.0072) | -0.0244 ***<br>(0.0072) | -0.0175 **<br>(0.0059) |
| randomxEDAY | | | 0.0096<br>(0.0112) | 0.0096<br>(0.0112) | 0.0102<br>(0.0112) | 0.0096<br>(0.0112) | 0.0102<br>(0.0112) | 0.0106<br>(0.0105) |
| NESEDAY | | | | 0.1223 ***<br>(0.0097) | 0.1219 ***<br>(0.0097) | 0.0976 ***<br>(0.0105) | 0.0980 ***<br>(0.0104) | 0.0694 ***<br>(0.0094) |
| AGG_EDAY | | | | | -0.3441 ***<br>(0.0649) | | -0.3295 ***<br>(0.0611) | -0.2042 ***<br>(0.0595) |
| AGG_NESEDAY | | | | | | -0.2051 ***<br>(0.0519) | -0.1986 ***<br>(0.0507) | -0.1037 *<br>(0.0519) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Adjusted $R^2$* | 0.4754 | 0.4754 | | 0.4758 | 0.4759 | 0.4759 | 0.4759 | 0.5337 |
| *N (Sectors)* | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 | 2464 (11) |
| *n* | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 | 3,624,621 |
| *Date FE? (Sector FE?)* | YES | YES | YES | YES | YES | YES | YES | YES (YES) |
| *Other controls?* | YES | YES | YES | YES | YES | YES | YES | YES |

Standard errors are double clustered by both firm and date. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

# 6. Conclusions

In this paper, we have investigated the relationship between information consumption and price informativeness, as well as delved into the foundations of these two strands of literature. We contribute to the literature by using a behavioral approach in investigating price informativeness, as well as exploring the foundations of information consumption. This background is the basis for the research questions, which are set to seek the sign and characteristics of the relationship between information consumption and price informativeness. They are then supported by various robustness checks, to ensure as robust conclusions as possible, within the scope of this paper.

Our measure of information consumption has been largely based on Ben-Rephael et al. (2021), and our results indicate similar interpretations. Similar to Ben-Rephael et al. (2021), we find that the main drivers for abnormal institutional investor consumption are earnings announcements and non-earnings scheduled events, and we too, find that the effect of earnings announcements is of higher economic significance. In alliance with Ben-Rephael et al. (2021), we find an information spillover effect, as related firms engage in information events. In contrast to Ben-Rephael et al. (2021), we do not find an effect of statistical significance on institutional attention by macro announcements. This finding is not in fundamental conflict with the findings of Ben-Rephael et al. (2021), though, as they only find support for macro announcement as a determinant of institutional attention at the lowest possible significance, $p<0.05$. Due to omitted variable bias in our sample, we conclude that further research is warranted to form a meaning on whether macro announcement is a determinant of institutional attention.

We find a positive correlation between our proxy for institutional information consumption - *abnormal institutional attention* - and our proxy for price informativeness - *price non-synchronicity*. Our evidence suggests that institutional attention has both economic and statistically significant, positive effects on price informativeness. This finding is robust for two different proxies of price informativeness. We also perform a placebo test and find strong evidence that the effect is robust and derived from our proxy of institutional attention. Moreover, we find indications that the events that have the largest effects on price informativeness are the unanticipated events. Even after introducing a post-earnings announcement drift, to account for eventual inertia in price informativeness, we find robust evidence that indicates an effect on price non-synchronicity by abnormal institutional attention.

For retail investors, we find no economic or significant correlation between retail investor attention and price non-synchronicity, indicating that retail information consumption has little or no effect on price informativeness, and should not be considered a driver of price informativeness. This finding is in line with prior studies (Ben-Rephael et al., 2021, 2017).

Limits to our study are mainly in direct consequence of lack of access to data, such as the NYSE Trade and Quote (TAQ) database and RavenPack. This entails an omitted robustness check in the shape of a triangulation effort through a third proxy for price informativeness; as well as an omitted variable bias in our evidence regarding information consumption. Our findings are furthermore vulnerable to possible reverse causality, possibly rendering them misleading. This warrants further research on the foundations of this paper.

# 7. References


Alin, A., 2010. Multicollinearity. *WIREs Comp Stat*, 2: 370-374.

An, H. and Zhang, T., 2013. Stock price synchronicity, crash risk, and institutional investors, *Journal of Corporate Finance*, Volume 21, Pages 1-15.

Ball, R., and Brown, P., 1968. An empirical evaluation of accounting numbers. *Journal of Accounting Research* 6: 159–78.

Barber, M.B. and Odean, T., 2008,. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *The Review of Financial Studies*, Volume 21, Issue 2, Pages 785–818.

Basak, S. and Pavlova, A.,. 2013. "Asset Prices and Institutional Investors." *American Economic Review*,103 (5): 1728-58.

Ben-Rephael, A., Da, Z, Israelsen, D.R,. 2017. It Depends on Where You Search: Institutional Investor Attention and Underreaction to News, *The Review of Financial Studies*, Volume 30, Issue 9, Pages 3009–3047

Ben-Rephael, Carlin, Da, and Israelsen., 2021. "Information Consumption and Asset Pricing." *The Journal of finance* (New York) 76.1: 357–394.

Bond, P., Edmans, A. and Goldstein, I., 2012. The Real Effects of Financial Markets. *Annual Review of Financial Economics,* 4(1), pp. 339-360.

Bernard, V.L., and Thomas, J.K., 1989. "Post-earnings-announcement drift: delayed price response or risk premium?." *Journal of Accounting research* 27: 1-36.

Beuselinck, C., et al., 2010. Mandatory IFRS reporting and stock price informativeness. *Center for Economic Research.*

Bhushan, R., 1989. Firm characteristics and analyst following, *Journal of Accounting and Economics*, Volume 11, Issues 2–3, Pages 255-274.

Bonsall,B.S., Green, J., and Muller, K.A., 2020. Market uncertainty and the importance of media coverage at earnings announcements, *Journal of Accounting and Economics*, Volume 69, Issue 1, 101264.

Brickley, J.A., 1986. Interpreting Common Stock Returns around Proxy Statement Disclosures and Annual Shareholder Meetings."*The Journal of Financial and Quantitative Analysis*, vol. 21, no. 3, pp. 343–349.

Chae J, Kim R, and Han J., 2020. Investor Attention from Internet Search Volume and Underreaction to Earnings Announcements in Korea. *Sustainability*. 12(22):9358.

Chen, Q., Goldstein, I., and Jiang, W., 2007. Price Informativeness and Investment Sensitivity to Stock Price, *The Review of Financial Studies*, Volume 20, Issue 3, Pages 619–650

Condie, S., Ganguli, J., & Illeditsch, P. K., 2013. Information Inertia. Working paper. Retrieved from: https://repository.upenn.edu/fnce_papers/25

Da, Z., Engelberg, J. and Gao, P., 2011. In Search of Attention. *The Journal of Finance*, 66: 1461-1499.

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

Dellavigna, S. and Pollet, J.M., 2009. Investor Inattention and Friday Earnings Announcements. *The Journal of Finance*, 64: 709-749.

Durnev, A., Morck, R., Yeung, B. and Zarowin, P., 2003 Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing?. *Journal of Accounting Research*, 41: 797-836.

Durnev, A., Li, K., Mørck, R. and Yeung, B., 2004. Capital markets and capital allocation: Implications for economies in transition*. *Economics of Transition*, 12: 593-634.

Engelberg, J.E. and Parsonos, C.A., 2011. The Causal Impact of Media in Financial Markets. *The Journal of Finance*, 66: 67-97.

Engelberg, J., Sasseville, C. and Williams, J., 2012. Market Madness? The Case of Mad Money. *Management Science,* 58(2), pp. 351-364.

Fama, E.F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, vol. 25, no. 2, 1970, pp. 383–417.

Ferreira, D., Ferreira, M.A., and Raposo, C.C, 2011. Board structure and price informativeness, *Journal of Financial Economics*, Volume 99, Issue 3, Pages 523-545.

Fort, T., Haltiwanger, J., Jarmin, R. et al., 2013. How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size. *IMF Econ Rev* 61, 520–559.

Foster, G., et al., 1984. "Earnings Releases, Anomalies, and the Behavior of Security Returns." *The Accounting Review*, vol. 59, no. 4, pp. 574–603.

Gabaix, X., et al., 2006. Institutional Investors and Stock Market Volatility, *The Quarterly Journal of Economics*, Volume 121, Issue 2, Pages 461–504.

Gelb, D.S., Zarowin, P., 2002. Corporate Disclosure Policy and the Informativeness of Stock Prices. *Review of Accounting Studies* 7, 33–52.

Gyntelberg, J., et al., 2009., Private Information, Stock Markets, and Exchange Rates (February 1, 2009). BIS Working Paper No. 271.

Hou, K., Xiong, W., Peng, L., 2009. A tale of two anomalies: The implications of investor attention for price and earnings momentum. Working paper, available at SSRN 976394.

Illeditsch, P.K., 2011. Ambiguous Information, Portfolio Inertia, and Excess Volatility. *The Journal of Finance*, 66: 2213-2247.

Jin, L., and Myers, S.C., 2006. R2 around the world: New theory and new tests, *Journal of Financial Economics*, Volume 79, Issue 2, Pages 257-292.

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

Kim, J.H., 2019.. "Multicollinearity and misleading statistical results." *Korean journal of anesthesiology* vol. 72,6: 558-569.

Liang, L., 2003. Post-Earnings Announcement Drift and Market Participants' Information Processing Biases. *Review of Accounting Studies* 8, 321–345.

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*, Volume 58, Issues 1–2, Pages 215-260.

Newhard, J.M, 2014.The stock market speaks: How Dr. Alchian learned to build the bomb, *Journal of Corporate Finance*, Volume 27, Pages 116-132,

Roll, R., 1984. Orange Juice and Weather. *The American Economic Review, 74*(5), 861-880.

Roll, R., 1988. R2. *The Journal of Finance*, 43: 541-566.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics*, Volume 80, Issue 2, Pages 309-349.

Sanders, R., & Zdanowicz, J., 1992. Target Firm Abnormal Returns and Trading Volume Around the Initiation of Change in Control Transactions. *The Journal of Financial and Quantitative Analysis,27*(1), 109-129.

Savor, P. and Wilson, M., 2016. Earnings Announcements and Systematic Risk. *The Journal of Finance*, 71: 83-138.

Stoll, H.R, and Whaley, R.E., 1990. Stock Market Structure and Volatility, *The Review of Financial Studies*, Volume 3, Issue 1, Pages 37–71.

Wurgler, J., 2000. Financial markets and the allocation of capital, *Journal of Financial Economics*, Volume 58, Issues 1–2, Pages 187-214.

# Appendix A. Variable Definitions

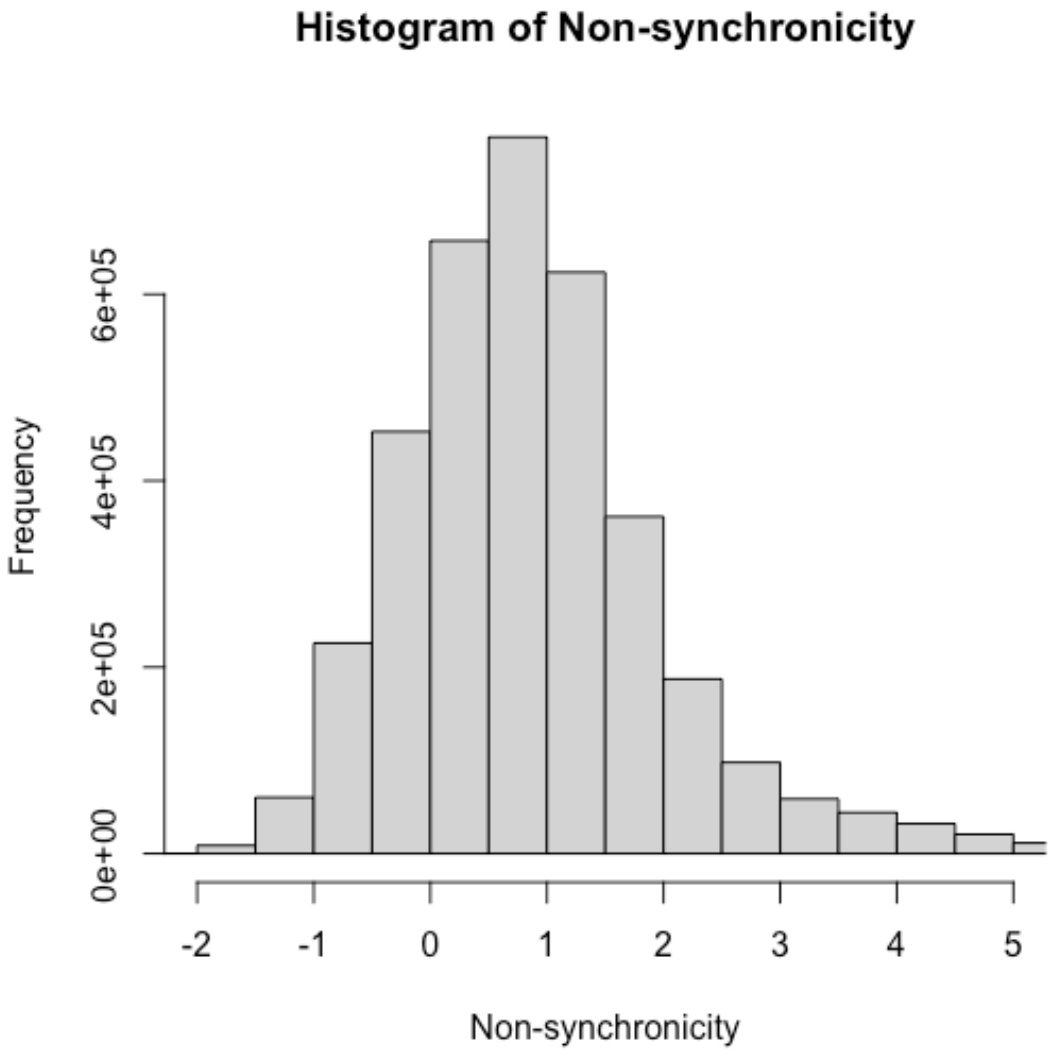
## Table A.1 Variable Definitions

| Variable | Definition |
|---|---|
| *Dependent variables* | |
| NS | Non-synchronicity, calculated daily, based on the $R^2$ from the Fama-French three factor model with data from the previous 252 trading days. Defined as the logarithmic transformation of: $(1\text{-}R^2)/\,R^2$. A higher NS indicates higher price informativeness, since less of the firm-specific returns can be explained by the Fama-French three factor model. Data for this variable is obtained from CRSP. |
| ILLIQ | Amihud (2002)'s illiquidity ratio. Defined as the average of the past 10 days' absolute returns divided by the dollar trading volume, DVOL. A higher illiquidity ratio indicates higher price informativeness. Data for this variable is obtained from CRSP. |
| AIA | Bloomberg records the number of times the news articles on a particular stock are read by its terminal users and the number of times the users actively search for news for a particular stock. Bloomberg then assigns a score of one to each article read and 10 for each news search. These scores are aggregated by Bloomberg into hourly counts. Using the hourly counts, Bloomberg creates a numerical attention score for each stock by comparing the past eight-hour average count to all hourly counts the past month for the same stock. Then, they assign a value of 1, 2, 3, or 4 if the average is between 80% and 90%, 90% and 94%, 94% and 96%, or greater than 96% of the previous 30 days' hourly counts, respectively. Finally, Bloomberg aggregates up to the daily frequency by taking a maximum of all hourly scores throughout the day. We are interested in abnormal attention, and thus we create AIA, defined as to be equal to one when Bloomberg's score is 3 or 4, and zero otherwise. |
| *Independent variables* | |
| ARA | A dummy variable created in a similar way to AIA. Based on Google Search Volume data, we create a measure called Abnormal Daily Search Volume Index (ADSVI), which is defined as the logarithmic transformation of the daily search volume index for firm *i* divided by the monthly average daily search volumes. Then, we assign a value to each ADSVI-value based on how it compares to the previous month's ADSVIs. As in the creation of AIA, if the ADSVI is larger than 94% of the previous month's values, we assign ARA a value of one. Otherwise the variable is zero. This is intended to capture abnormal attention from retail investors. |
| random | A random dummy variable, created by a binomial distribution with a 0.0765 probability of being equal to one. This is done to create a random variable that is similar to AIA. |
| EDAY | A dummy variable that is equal to one on earnings announcement days for firm i and zero otherwise. Earnings announcements data are from I/B/E/S. |

| | |
|---|---|
| NESEDAY | A dummy variable that is equal to one on days with non-earnings scheduled events and zero otherwise. Corporate events data are from Bloomberg's event calendar, (Bloomberg command "EVTS"). |
| AGG_EDAY | A value-weighted average of all, except firm i, earnings announcing firms in the sample for a particular day. The value weights come from the combined market values in the sample. |
| AGG_NESEDAY | A value-weighted average of all, except firm i, firms with non-earnings scheduled events for a particular day. The value weights come from the combined market values in the sample. |
| GDP | A dummy variable equal to one on days with an announcement of the "advance" estimate of quarterly U.S. Gross Domestic Product by the Bureau of Economic Analysis, and zero otherwise. Announcement dates come from Bloomberg. |
| NFP | A dummy variable equal to one on days with an announcement of the U.S. nonfarm payroll statistics by the Department of Labor, and zero otherwise. Announcement dates come from Bloomberg. |
| PPI | A dummy variable equal to one on days with an announcement of the U.S. Producer Price Index numbers by the Bureau of Labor Statistics, and zero otherwise. Announcement dates come from Bloomberg. |
| FOMC | A dummy variable equal to one on days with an announcement of the Federal Open Market Committee rate decision, and zero otherwise. Announcement dates come from Bloomberg. |
| ISM | A dummy variable equal to one on days with an announcement of the Institute for Supply Management Manufacturing statistics by Bureau of Labor Statistics, and zero otherwise. Announcement dates come from Bloomberg. |
| MACRO | A dummy variable equal to one if at least one of NFP, PPI, FOMC, GDP, and ISM is equal to one, and zero otherwise. |
| Leverage | The firm's leverage, calculated as the ratio between long-term debt and total assets, using Compustat data. |
| DVOL | The firm's daily dollar trading volume in millions of dollars. |
| abs_ret | The absolute value of the firm's daily total return. Data for daily returns are from CRSP. |
| b_mkt | The firm's CAPM beta, calculated for each day based on the previous 252 trading days. |
| lnSize | The natural logarithm of the stock's market capitalization. |
| LnBM | The natural logarithm of the firm's book-to-market ratio. |
| insthold | The percentage of shares held by institutional investors. Data on institutional ownership are from Bloomberg. |

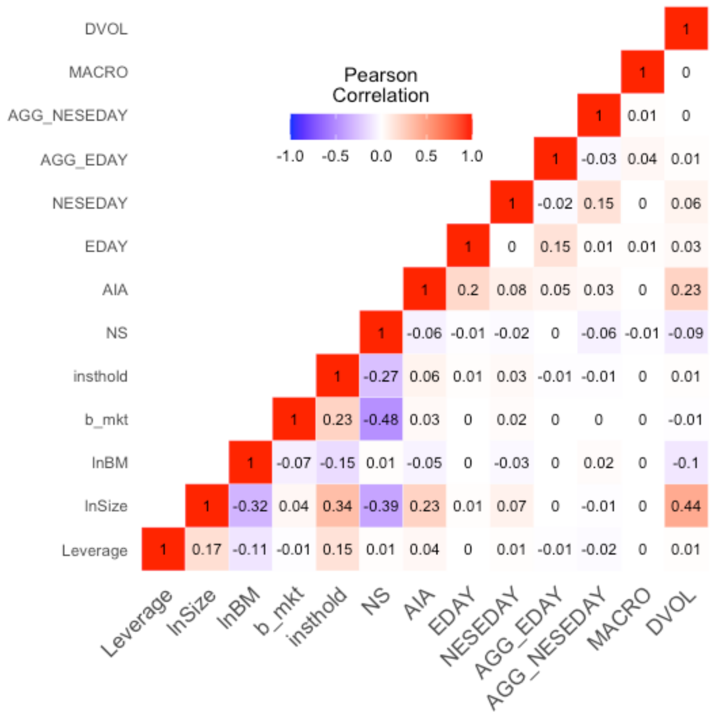
# Appendix B. Histogram of Non-Synchronicity

**Figure B.1 Histogram of Non-Synchronicity**



# Appendix C. Correlation Matrix

**Figure C.1 Correlation matrix**




# Appendix D. Multicollinearity tests for Table II

A variance inflation factor is an indicator of multicollinearity in OLS-models. It is calculated for each explanatory variable by running a regression of the explanatory variable Xi on all other explanatory variables, the coefficient of determination, Rsq, is then used to calculate the VIF. (Alin, 2010) Traditionally, values larger than 10, some also use 5, indicate problems with multicollinearity. In the table below, we see that all VIF-values are just above 1, which implies that we have no serious problem with multicollinearity.

**Table D.1 Variance Inflation Factors for Model 6 in Table II.**

| Variables | Tolerance | Variation Inflation Factor |
|---|---|---|
| EDAY | 0.9692 | 1.0318 |
| NESEDAY | 0.9722 | 1.0286 |
| MACRO | 0.9853 | 1.0149 |

| | | |
|---|---|---|
| AGG_EDAY | 0.9203 | 1.0866 |
| AGG_NESEDAY | 0.7934 | 1.2603 |
| lnBM | 0.9628 | 1.0387 |
| abs_ret | 0.9652 | 1.0360 |
| DVOL | 0.9858 | 1.0144 |
| Leverage | 0.9752 | 1.0255 |
| b_mkt | 0.9705 | 1.0304 |

Another method to investigate the potential of multicollinearity is to calculate the eigenvalues from a matrix composed of standardized explanatory variables. From the eigenvalues, it is possible to calculate a condition index. The largest condition index is the condition number, bold in the table below, and a condition number larger than 30 indicates strong multicollinearity. (Kim, 2019)

**Table D.2 Eigenvalues and Condition index for Model 6 in Table II**

| Eigenvalue | Condition Index | EDAY | NESEDAY | MACRO | AGG_EDAY | AGG_NESEDAY | lnBM | abs_ret | DVOL | Leverage | b_mkt |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.5524 | 1.0000 | 0.0015 | 0.0015 | 0.0065 | 0.0082 | 0.0073 | 0.0092 | 0.0090 | 0.0031 | 0.0090 | 0.0039 |
| 1.1586 | 2.1891 | 0.0003 | 0.1246 | 0.0148 | 0.0024 | 0.0339 | 0.0042 | 0.0017 | 0.0007 | 0.0043 | 0.0006 |
| 1.0714 | 2.2765 | 0.2962 | 0.0617 | 0.0012 | 0.0585 | 0.0039 | 0.0024 | 0.0008 | 0.0027 | 0.0017 | 0.0002 |
| 1.0163 | 2.3374 | 0.0135 | 0.0000 | 0.0433 | 0.0001 | 0.0001 | 0.0013 | 0.0001 | 0.0030 | 0.0009 | 0.0001 |
| 0.9743 | 2.3872 | 0.0521 | 0.3392 | 0.0035 | 0.0053 | 0.0022 | 0.0020 | 0.0013 | 0.2355 | 0.0001 | 0.0001 |
| 0.9648 | 2.3989 | 0.4370 | 0.0280 | 0.0081 | 0.0008 | 0.0034 | 0.0000 | 0.0018 | 0.2370 | 0.0026 | 0.0004 |
| 0.8696 | 2.5268 | 0.0740 | 0.3972 | 0.0078 | 0.0004 | 0.0006 | 0.0039 | 0.0034 | 0.4636 | 0.0000 | 0.0001 |
| 0.7678 | 2.6891 | 0.0002 | 0.0015 | 0.8348 | 0.0026 | 0.0001 | 0.0268 | 0.0090 | 0.0276 | 0.0309 | 0.0014 |
| 0.6193 | 2.9942 | 0.0124 | 0.0050 | 0.0048 | 0.0422 | 0.0057 | 0.1276 | 0.5786 | 0.0035 | 0.1301 | 0.0021 |
| 0.6108 | 3.0150 | 0.0901 | 0.0169 | 0.0352 | 0.6980 | 0.0130 | 0.0537 | 0.0019 | 0.0029 | 0.0594 | 0.0000 |
| 0.4692 | 3.4400 | 0.0104 | 0.0105 | 0.0227 | 0.0273 | 0.1944 | 0.4585 | 0.2015 | 0.0114 | 0.0078 | 0.0030 |
| 0.4562 | 3.4889 | 0.0045 | 0.0063 | 0.0065 | 0.0104 | 0.0708 | 0.2337 | 0.0750 | 0.0065 | 0.6737 | 0.0032 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2561 | 4.6563 | 0.0036 | 0.0068 | 0.0051 | 0.1154 | 0.6067 | 0.0568 | 0.0860 | 0.0000 | 0.0424 | 0.1222 |
| 0.1590 | 5.9089 | 0.0000 | 0.0000 | 0.0038 | 0.02134 | 0.0285 | 0.0071 | 0.0276 | 0.0003 | 0.0003 | 0.3464 |
| **0.0541** | **10.1233** | 0.0000 | 0.0008 | 0.0019 | 0.0072 | 0.0300 | 0.0128 | 0.0025 | 0.0021 | 0.0367 | 0.5161 |

# Appendix E. Multicollinearity tests for Table III

Similar to Appendix D, we calculate the variation inflation factors for the explanatory variables in Model 6 in Table III. Unfortunately, we found no, simple enough, way to calculate the Eigenvalues and Condition Index for panel models. The results are seen in the table below and they indicate that multicollinearity does not pose a large problem in our analysis.

**Table E.1 Variance Inflation Factor for Model 6 in Table III.**

| Variable | Tolerance | Variance Inflation Factor |
|---|---|---|
| AIA | 0.8680 | 1.1520 |
| EDAY | 0.6431 | 1.5550 |
| AIAxEDAY | 0.6232 | 1.6047 |
| NESEDAY | 0.8617 | 1.1605 |
| AGG_EDAY | 0.8808 | 1.1354 |
| AGG_NESEDAY | 0.8476 | 1.1798 |
| Leverage | 0.9535 | 1.0487 |
| insthold | 0.8197 | 1.2199 |
| lnSize | 0.7487 | 1.3356 |
| lnBM | 0.8967 | 1.1152 |
| b_mkt | 0.9345 | 1.0700 |