

**UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS  
ACADEMIC RESEARCH AND GRADUATE UNIT  
GRADUATE PROGRAM IN ACCOUNTING  
MASTER'S DEGREE**

**MATHEUS RAMOS DE CASTRO GONZALEZ**

**INFORMATION ASYMMETRY INDEX:  
The View of Market Analysts**

**Porto Alegre**

**2019**

MATHEUS RAMOS DE CASTRO GONZALEZ

**INFORMATION ASYMMETRY INDEX:  
The View of Market Analysts**

Thesis presented as a partial requirement  
to obtain a Master's Degree in Accounting,  
by the Graduate Program in Accounting of  
Universidade do Vale do Rio dos Sinos -  
UNISINOS

Advisor: Prof. Dr. Roberto Frota Decourt

Porto Alegre

2019

G643i Gonzalez, Matheus Ramos de Castro.

Information asymmetry index : the view of market analysts / Matheus Ramos de Castro Gonzalez. – 2019.

60 f. : il. ; 30 cm.

Dissertação (mestre) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Contabilidade, 2019.

“Advisor: Prof. Dr. Roberto Frota Decourt.”

Dados Internacionais de Catalogação na Publicação (CIP)  
(Bibliotecária: Amanda Schuster – CRB 10/2517)

MATHEUS RAMOS DE CASTRO GONZALEZ

**INFORMATION ASYMMETRY INDEX:  
The View of Market Analysts**

Thesis presented as a partial requirement  
to obtain a Master's Degree in Accounting,  
by the Graduate Program in Accounting of  
Universidade do Vale do Rio dos Sinos -  
UNISINOS

EXAMING BOARD

---

Prof. PhD. Heitor Almeida – ILLINOIS

---

Prof. PhD. Cristiano Machado Costa – UNISINOS

---

Prof. Dr. João Zani – UNISINOS

---

Advisor - Prof. Dr. Roberto Frota Decourt – UNISINOS

## **ACKNOWLEDGMENT**

I wish to thank, first and foremost, my thesis advisor Dr. Roberto Decourt for his excellence on steered me in the right direction. He was always available to share his knowledge when I was in difficult situation, without him this work wouldn't been possible. I would also like to thank the thesis committee for the contribution they've made, surely was very helpful to enhance the validation of this research.

In addition, a special thanks for Dr. Clea Beatriz Macagnan, Dr. Daniel Vancin, Md. Carlos Salami, Md. Thobias Bassotto Zani, BM. Cacioano Gianechini, Econ. Carlos Damião Prolo Junior and Econ. Louise Salami for their advises and support. I am gratefully indebted for their valuable contributions on this thesis.

Finally, I must be grateful to my family, friends and classmates for providing me the support needed on the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

## **ABSTRACT**

Information Asymmetry impact companies over different format, developing a great interest of researchers to study this construct. However, given those multiple impacts, distinctive proxies haven been used, quantitative and qualitative, but so far haven't found appropriate model to measure the information asymmetry degree. This work aims to fill this gap proposing a model with secondary data capable to measure information asymmetry through market perspective. In this context, initially a survey was conducted between 2016 and 2018 with market analyst accredited by CFA, or CNPI. Based on their opinion an index was built in order to ranking the companies by their perception of better to worst disclosure. To validate the index, it was tested in a pooled cross-section model with dummies for sector and time for three groups of proxies: External Analysis, Internal Analysis and Market Microstructure, designed by the source of the proxies. The findings show that Volatility, Growth Opportunities and Coverage plays an important role in the way to determine companies' information asymmetry degree. At the end, this work proposed a model for future researcher on the field.

**Key Words:** Information Asymmetry. Information Asymmetry Index. Disclosure. Sell Side Coverage. Market Microstructure.

## RESUMO

Assimetria Informacional impacta empresas sobre diferentes formatos, formando um grande interesse por parte dos pesquisados a estudarem este constructo. No entanto, dado aos seus múltiplos impactos, diferentes proxies são usadas, sendo estas quantitativas e qualitativas, mas até o momento não foi achando um modelo único apropriado para mensurar níveis de assimetria informacional. Este trabalho objetiva preencher esta lacuna, propondo um modelo com dados secundários capaz de medir assimetria informacionais sob a ótica de mercado. Neste sentido, inicialmente foi conduzida uma *survey* entre 2016 e 2018, com analistas de investimentos cujos, possuem certificação CFA, ou CNPI. Baseado em suas respostas foi criado um índice com o objetivo de ranquear as empresas entre as que possuem melhor e pior níveis de transparência. Para validar este índice, foi testado através de um modelo de corte transacional com dados empilhados e com *dummies* para setor e tempo, para três grupos de proxies: Análise Externa, Análise Interna e Microestrutura de Mercado, desenhados a partir da origem das proxies. OS resultados mostram que Volatilidade, Oportunidades de Crescimento e o Nível de Cobertura por Analistas de Mercado têm um papel importante para determinar o nível de assimetria informacional das empresas. Ao final do trabalho é proposto um modelo a ser utilizado em pesquisas futuras na área.

**Palavras Chave:** Assimetria Informacional. Índice de Assimetria Informacional. Disclosure. Cobertura de Sell Side. Microestrutura de Mercado

## LIST OF CHARTS

Graph 1 - K-Factor used in Elo Rating .....	24
---	----



## LIST OF FIGURES

Figure 1 - Tree diagram of the trading process .....	31
--	----

## LIST OF TABLES

Table 1 - Difference in Elo-rating and the corresponding win expectation.....	24
Table 2 - External Analysis Proxies.....	27
Table 3 - Internal Analysis Proxy.....	29
Tabela 4 - Market Microstructure Proxy .....	33
Table 5 - Control Variable Proxies.....	34
Table 6 - Descriptive Statistic IAI.....	36
Table 7 - Descriptive Statistic.....	37
Table 8 - Matrix Correlation.....	38
Table 9 - Polled Cross Section – All Sample 2016-2018.....	39
Table 10- Polled Cross Section – All Sample 2016-2018.....	41
Table 11- Robustness test ADR inclusion.....	43
Table 12 – Robustness Test PIN inclusion.....	44
Table 13- PIN Descriptive Statistic.....	48

## LIST OF ACRONYMS

CVM	Brazilian Security Commission
APIMEC	Analysts and Investment Professionals of Capital Markets Association
CFA	Chartered Financial Analyst
IAI	Informational Asymmetry Index
R&D	Research and Development
IBrA	Brazil Broad-Based Index
CAPM	Capital Asset Pricing Management
B3	Brazilian Stock Market
BACEN	Brazilian Central Bank

## TABLE OF CONTENTS

<b>1 INTRODUCTION.....</b>	<b>13</b>
1.1 THEME.....	15
1.2 DELIMITATION OF THE STUDY .....	15
1.3 RESEARCH PROBLEM.....	15
1.4 OBJECTIVE .....	16
1.5 JUSTIFICATION.....	16
<b>2. LITERATURE REVIEW.....</b>	<b>18</b>
2.1 INFORMATION ASYMMETRY.....	18
2.2. ASYMMETRIC INFORMATION PROXY'S REVIEW.....	21
<b>3 METHODOLOGY.....</b>	<b>23</b>
3.1 DEPENDENT VARIABLE.....	23
3.2 INDEPENDENT VARIABLE .....	26
3.1.1 EXTERNAL ANALYSIS PROXIES.....	26
3.1.2 INTERNAL ANALYSIS PROXY.....	28
3.1.3 MARKET MICROSTRUCTURE PROXY .....	29
3.3 CONTROL VARIABLE .....	33
3.4 ECONOMETRIC MODEL.....	34
<b>4 OUTCOME ANALYSIS.....</b>	<b>36</b>
4.1 STATISTIC DESCRIPTIVE .....	36
4.2 CORRELATION MATRIX.....	38
4.3 OUTCOMES PRESENTATION.....	39
4.4 ROBUSTNESS TEST.....	43
<b>5 DISCUSSION.....</b>	<b>46</b>
5.1 EXTERNAL ANALYSIS OUTCOMES .....	46
5.2 INTERNAL ANALYSIS OUTCOME.....	47
5.3 MARKET MICROSTRUCTURE OUTCOME .....	48
5.4 DUMMIES OUTCOME .....	49
5.5 GENERAL FINDINGS .....	50
<b>6. CONCLUDING REMARKS.....</b>	<b>52</b>
<b>REFERENCES.....</b>	<b>54</b>
<b>APPENDIX A – ECONOMETRIC TESTS .....</b>	<b>58</b>
A.1 MULTICOLLINEARITY TEST – VARIANCE INFLATION TEST .....	58

A.2 HOMOSCEDASTICITY TESTS .....	58
<b>APPENDIX B – DISCLOSURE WEBSITE.....</b>	<b>60</b>
<b>APPENDIX C – PLOTS .....</b>	<b>61</b>

## 1 INTRODUCTION

Information asymmetry can be perceived over different forms, many researches have been dedicated to study each of these different impacts on companies. Verrechia (1983), Myers and Majluf (1985), Easley and O'Hara (1987), Harris and Raviv (1991) and Wang (1993) show theoretically that opaque companies tend to have more noise, volatility and informed traders acting on pricing, as well as more conflict between managerial and shareholders which impacts on the firm's investment decision and capital cost. Under those circumstances, information asymmetry has been receiving much attention of modern literature.

As seen in Beyer et al (2010) many academics have been developing their researches using different proxies, quantitative and qualitative variables, looking to verify the impact of information asymmetry into the companies in terms of executive compensation, cost of capital, level of indebtedness, profitability, shareholder return, liquidity, control structure dividend policy, asset pricing and others, but so far haven't found appropriate model to measure the information asymmetry degree. Most corporate finance literature have been using proxies from financial analysts forecast, company's investment opportunities, and the presence of informed and informed traders over daily stock prices.

This work conducted a survey with financial analysts in order to verify their perception of disclosure and information asymmetry over companies which belong to Brazil Broad-Based Index (IBrA). A financial analyst was invited to access a web site and choose among a pair of companies which one had a better disclosure. The analysts must be certified by Analysts and Investment Professionals of Capital Markets Association of Brazil (*APIMEC*), or Charter Financial Analysts (CFA) holder.

Based on Elo (1961) algorithm this work creates the Information Asymmetry Index (IAI) classifying Brazilian companies' disclosure under market perception. The logic of this ranking is to check the likelihood of a win (loss) between direct disputes. A win when the expected probability is high will add very few points to the ranking, however, a win with a very small probability adds many points to the ranking. On the other hand, a loss with high probability loses a few points and, a loss with low probability loses many points.

A concern of this work is if the IAI correctly captures the disclosure perception of a financial analyst, or if it is disturbed by other sources besides disclosure. In order

to test the robustness of IAI correct, three groups of proxies were built given literature review. The categories are classified according to proxies' sources: external analysis, internal analysis and market microstructure.

External analysis is a group characterized by information that came from out of the company, i.e. financial analysts and stock price volatility. They are qualified professionals who work to understand every companies' details with the goal of determining what should the stock price be for a specific period. Also, the path to correct stock price bring volatility to stock if companies doesn't provide enough disclosure for investors, in this sense volatility can be a measured of how investors perceived companies' opaqueness. Analyst coverage, the difference between the actual earnings per share and the Bloomberg median forecast and stock price volatility will be use as information asymmetry proxies in this group.

Internal analysis looks for information asymmetry trough companies' financial reports and stock price volatility. Shin and Stulz (2000) studied the relation between Tobins'Q and systematic equity risk and total equity risk. They find that firms with higher market to book ratio have higher growth opportunities. High levels of market to book ratio means that investors are willing to buy shares for a more expansive than book value, this can possible because company disclosure higher future growth. If this are correct, market-to-book ratio is associated to a lower levels of information asymmetry. This work will use growth opportunities as proxy for this group.

Market microstructure investigated the likelihood of informed investors determined stock price. Easley, Kiefer and O'Hara (1997), Easley, Hvidkjaer, and O'Hara (2002), Easley and O'Hara (2004) and Hvidkjaer (2010) developed a model to demonstrate that informed investors not only play a critical component on stock price formation, but also build a different portfolio from other investors, which is the opposite of what the CAPM theory sustained. In addition, the probability of informed traders (PIN) has no bias, since is no under disclosure regulation, or researcher interpretation, it simply follows the buy or sell quantity orders. This work will use the probability of informed trading as proxy for this group.

## **1.1 Theme**

Measures of Information Asymmetry on public held companies.

## **1.2 Delimitation of The Study**

The present research is going to review the recent and relevant works of information asymmetry specifically on which are the proxies and the methodologies used to measure this construct in order to face IAI ranking to test its efficacy.

The survey was conducted only with certified financial analysts. Besides, it is possible to assume that there are some good analysts working in the market who are not certified, the work focuses only on those certified ones to enhance robustness.

The sample will be around 120 publicly-held companies from the Brazil Broad-Based Index (IBrA), traded at B3, the Brazilian stock market. The number of the companies, and the companies can vary over time, given IBrA methodology

## **1.3 Research Problem**

The finance theory's premise that companies and their management work to maximize value to their stockholders, i.e. get the company value as high as possible. In this sense, the best practices that could lead the company on the way of transparency and disclosure actions emerges from the corporate governance, achieving a lower degree of risk, so that the market starts trusting the company and their executives. Hence, companies should strive to a higher level of disclosure, even though that would be an additional cost involved.

In effort to measure the impact of information asymmetry on companies, several studies have been conducted, because the level of information asymmetry between company managers and the market may have different consequences for each company, such as executive compensation, cost of capital, level of indebtedness, profitability, shareholder return, liquidity, governance and dividend policy, consequently managers are constantly in a trade-off about what information will be disclosed by the companies (BEYER, et al. 2010).

Because information asymmetry has a large impact on companies, there are plenty of methodologies to evaluate this construct likewise disclosure and quality



reports (BEYER, et al. 2010). Researchers bias their analyzes by perceiving information asymmetry only through the eyes of their work, but not in all impacts it may have on firms. Academics who observe information asymmetry in initial public offering, usually use growth opportunities as a proxy, living aside important details like how competitive the market is, how many hours of meeting did the company have before it goes public and if the company already public in other country, all of it, and more, may increase (decrease) the information asymmetry of the entrance company. Another strand of researchers limits its construct on firm's expected future earnings and market forecast, although it might be a relevant proxy, there isn't relevant news every day, while companies' stock price can have large volatility even in the days with absence news. For the same reason the presence of uninformed traders can dramatically change over time, not only by companies' news, but by changes in economy, survivorship bias and other behavioral issues and size of the market diminishing. All things considered, the absence of a methodology capable to capture all the aspects of information asymmetry reduces external validity of researchers. That is the central concern of this work, to provide a wide and highly accurate method to measure information asymmetry.

#### **1.4 Objective**

The purpose of this research is to create a model with secondary data reflecting the analysts' perception over companies' disclosure. This model will be provided starting with the Information Asymmetry Index and testing its robustness using methods of measure information asymmetry provided by the literature.

#### **1.5 Justification**

As seen in Beyer et al (2010) many academics have been developing their researches using different proxies, quantitative and qualitative variables, looking to verify the impact of information asymmetry into the companies in terms of executive compensation, cost of capital, level of indebtedness, the company's profitability, shareholder return, liquidity, control structure and dividend policy, but until now, hasn't been found an effective way to measure it.

Although there are plenty of researches over information asymmetry, the variability of the methodology used contributes negatively to the literature because it decreases the research's external validity. To my knowledge, few researches attempt to build a wide methodology to cover all aspects of information asymmetry and no one is based on market participant perception adherent on literature. Thus, this work aims to create an index under financial analyst perception and to test it over and test it with what the literature indicates as proxies of informational asymmetry.

Investment analysts have extensive work to do, to determine companies' value they need to fully understand its business, read their financial statement (including footnotes and some accessories commentary). Moreover, they are influenced by the cost of achieving information and more importantly, their capacity to prove their right instead of the market (BRENNAN & TOMAROWSKY, 2000). Hence, investment analysts are too deep in companies' day by day, figures and disclosures practices. In addition, investment analysts are concerned about liquidity which measures the investor's demand of a stock, Kyle (1985) said that as higher is the number of shares traded, the less would be the degree of asymmetry information. In other words, Bushe and Miller (2012) stated that firms with low visibility and poor disclosure programs move away from security analysts and institutional investors. Hence, companies with disclosure policy can enhance their liquidity (BOTOSAN, 1997) which can lead the market to a more accurate pricing. Since investment analysts contribute to enhancing capital markets through their corporate reports, valuations and forecasts (HEALY & PALEPU, 2001), it is plausible to assume that they are one of the most qualify agents to evaluate companies' disclosure, which consequently can decrease asymmetry information (DIAMOND & VERRECCHIA, 1991). This work chooses to survey the opinion of certified analysts by CFA institute or APIMEC, besides a good analyst could not hold one of its titles, the ones who hold it certainly possess the knowledge to conduct a great valuation and interpretation of companies' figures. Besides that, both institutes helped the research by asking their affiliates to answer the survey.

## 2. LITERATURE REVIEW

An important asset in finance and economy is information. Stiglitz (2017) stated that about a century ago economists started to study information economics, developing models carrying out the presumption of market efficiency aiming to understand the economic policy impacts. These studies revealed that quite often markets aren't efficient, which means that, consequently, information plays an important driver in the efficient capital allocation.

Information and knowledge are substantially different from ordinary good studied by economists, due its global and public characteristics (STIGLITZ, 1999). Verrechia (1983) defined it as a signal which reveals the true liquidating value of a risky asset perturbed by some noise and Usategui (2000) complemented this definition by adding that the most valuable information is the one that solves the uncertainty of the decision maker. The value of information is a puzzle to complete, Usategui (2000) argued that an information provides the decision maker a higher expected return, consequently, the value of an information is the difference between expected return with an additional information vis-à-vis the expected return without it. In his words, agents would be willing to pay this entire difference. Even though its a plausible method of evaluating information, it doesn't seem coherent that agents are willing to pay all the extra earning for that information. In order to pay it, they would be incurring some additional costs, turning the expected profit at the same level of what it was before without possessing the information. A fraction of the extra earning would be more reasonable, but still facing the problem of which fraction would be fair.

### 2.1 Information Asymmetry

Fama (1969) said that an efficient market is one which security prices at all the times "fully reflects" plenty of available information. In fact, this term is so general, that makes it difficult to test it empirically. Initially, he states some conditions to market efficiency: information is costless, available to all and easily understandable. In brief, there are three empirical test categories depending on the information interest: weak, semi-strong and strong form. In the weak form, it was tested if the information interest was historical price, he found evidences that daily price changing were dependent, proving a serial correlation, but close to zero said that an overreaction, might be followed

by large price changes, although with unpredictable sign, showing that investors take a while to understand and evaluate the new information, even though he found out that the first day's announcement is unbiased. Testing semi-strong form is a format that stock prices fully reflected all public information supports the theory of efficient market, i.e. future dividend payments, split announcements, earnings announcements, new issues, or other information are, on average, fully reflected in the prices. Also, there is the strong form which prices reflect all available information, however, two important deviations have been found, and it was discovered that some highly influenced market agents have access to information before others, making profit with it, and some corporate insiders can have access to some exclusive information about their companies, but even their price deviation would permanently persist.

Although, many researchers have been criticizing this view, Brennan and Tamarowski (2000) say that the initial conditions for market efficiency are strongly wrong in practice. Managing a company is truly complex, they must be aware of external threats, internal conflicts and they often sell technically sophisticated products, which may impact on share's values and can lead the financial market to misprice it.

On the other hand, the studying of market efficiency and information asymmetry have been emerging in areas like accounting, finance, and corporate governance. Akerlof (1970), described the market of lemons, where informal guarantees and asymmetric information takes place, in other words, adding the construct "trust" into an economic model. He noticed that in the market of used cars in America, asymmetry information was inherent. Because buyers can't identify the difference between good and bad cars, which are traded at the same price, the sellers of good cars would be discouraged to offer their assets, since they wouldn't get the expected value for the car, but in fact, the value of a lemon car. This process, named Adverse Selection, was detected in other markets too. Usategui (2002) examined this practice between companies and banks, stating that companies might have their own resources needed to finance a project, although as they are risk-averse, they're going to take a loan in a bank which is risk averse too, but in a lower rate. Whether the bank knows the risk distribution of his credit portfolio, the interest rate charged in each project would be a value that represents the average risk of all the loans. Hence, companies with lower risks may finance their projects with internal funds. For this reason, banks would have creditors with higher risks. As in the market of lemons (AKERLOF, 1970) by the adverse selection, only companies with high credit risks are going get a bank loan, making the

market worse. So, it might be plausible to assume that this phenomenon can take place when a company goes public. Underwriters force managements to issue equities below their expected return (STOLL & CURLEY, 1970), if it is true, only companies which don't have internal funds would go public, turning market poorer, putting away good companies.

Because information is difficult to evaluate, in the context of corporate finance, literature has brought the concept that firm's insiders are well informed than market participants, some researchers have been dedicated to study what is called "conflict of interest", especially in the relations between equity holders and managers and between equity holders and debtholders.

Harris and Raviv (1991) in their seminal article, provided a review of what had been written so far about agency costs, asymmetric information, and other topics. Agency costs is the cost due to conflict of interest, Harris and Raviv (1991), said it takes place by two different relations: between shareholders and managers and between shareholders and debtholders. The first conflict arises when there isn't an alignment in the companies' corporate governance. Managers, who don't have shares, can prefer personal compensation and a higher leverage – besides higher profits –, decreasing free cash flow to equity and consequently not maximizing firm value. In this sense, managers would be benefited by a companies' profit. Consequently, equity holders can be conservative to select companies' investments, even if they have a profitable payout. The second conflict occurs in the relation between debtholders and equity holders, because the covenants contracts lead equity holders to invest sub-optimally, in a process named "asset substitution effect". Equity holders will capture the gain of an investment only if it yields a return bigger than the cost of debt, otherwise only the debtholders would be benefited. This relation is an incentive to equity holders to invest in risky projects, even if they decrease the equity value, aiming to get higher returns. In addition, Rabelo and Vasconcelos (2002) found a third conflict between minority shareholders and controlling in Brazil, they said that ownership is too concentrate, in structure called pyramids, which enhance the power on dominant shareholders, and do not see minority shareholders as partners.

Information Asymmetry also impacts on capital structure and level of indebtedness, (Modigliani and Miller, 1958; Ross, 1977; Myers and Majuf, 1984; Botosan, 1997). Harris and Raviv (1991) stated that internal sources are always preferred than external, to avoid stock price reaction. However, companies go public to

financing, which can implicate a negative reaction on stock price, because investors might conclude that internal sources and riskless debt wasn't enough, or wasn't there for the company, requiring higher returns. Moreover, debt issuance is a signal of asymmetry information. Harris and Raviv (1991) argue that managers know about firms' returns distribution, companies are expected to leverage (deleverage) if the current market is lower (higher) than futures. Since investors would expect higher returns if the debt level was increasing – as higher quality finance firms issue more debt and lower quality companies issue more equity to finance – the stock prices' reaction should be positive in response to debt issuing.

## **2.2. Asymmetric Information Proxy's Review**

Information Asymmetry has been receiving relevant attention on the body of corporate finance literature, even though, there isn't a consensus in how to proxy information risk, since it is not an observable construct, empiricist must rely on proxy variables. Clarke and Shastri (2001) divided in three general classes of proxies:

External analysis is the first, literature has been using analysts forecast of future earnings as proxy of information asymmetry, researchers find that as long as companies increase communication, more accurate stock prices would be, there would also be more analyst coverage, less dispersion on analyst forecast and consequently a reduction in asymmetry (LANG & LUNDHOLM, 1993; THOMAS, 2002, IRANI & KARAMANOU, 2003).

Internal analysis is the second group, it looks for proxy in order to identify growth opportunities, since companies with higher growth opportunities have a higher degree of information asymmetry (ADAM & GOYAL, 1999; SHIN & STULZ, 2000). Literature has been using R&D investments, market-to-book asset ratio and earnings-price.

Finally, several papers had payed attention on the adverse selection component of bid-and-ask spread, since market makers are trading with unidentified investors in a competitive environment, they are widening the spread to recover possible losses traded with informed investors (GLOSTEN & HARRIS, 1988). Literature (LAMBERT, LEUZ & VERRECHIA, 2008; ARMSTRONG, CORE & TAYLOR, 2011; HE, LEPONE & LEUG, 2013) also examines the relation between information asymmetry and cost of capital and equity, the findings suggest a positive relation between them, especially when markets are imperfect. By modelling stock liquidity and the frequency of bid and ask

spread is plausible to make inferences about the likelihood of informed trading. Easley and O'Hara (2004) proposed a theoretical model based on the assumption that equilibrium investors would hold different portfolios due to their capability to obtain information. Informed investors will optimally diversify, although uninformed will not hold a stock to maintain an excess of assets with "bad news" and a few with "good news". Using a market microstructure model Easley, Hvidkjaer, and O'Hara (2002) show empirically how to measure the private measure information which is the third analysis group of this work.

### 3 METHODOLOGY

This chapter presents the econometric approach used to evaluate companies' data in the way to determine the disclosure ranking. Also, it is presented the dependent and independent variables, as followed.

#### 3.1 Dependent Variable

The Informational Asymmetry Index (here an after IAI) was created to capture the analysts' perceptions about the level of company's disclosure and information asymmetry. IAI is based on Elo ratings, which was developed by Elo (1961) and is best known as the ranking system used to rank chess players.

The logic of this ranking is to check the likelihood of a win (loss) between direct disputes. A win when the expected probability was high would add very few points to the ranking, however, a win with a very small probability adds many points to the ranking. On the other side, a loss with high probability loses a few points and, a loss with low probability loses many points.

The IAI will use this method on all pair company dispute which was answered by market analysts accredited on APIMEC (Analysts and Investment Professionals of Capital Markets Association) or CFA holder (Chartered Financial Analyst). Hence the IAI was able to capture the disclosure of many Brazilian companies from the market perspective.

In order to exemplify this logic, let's assume a dispute among two companies, company X (Elo-rating score: 1,200) and company Y (Elo-rating score: 1,000). The difference between rankings is 200 points, which would represent a win probability of 76% for X and 24% for Y, according to table 1 presented, by Albers and Vries (2001). The new companies' score would be as followed:

Equation 1: IAI score

$$\text{New X Score} = \text{Previous Rank Score} + (1 - p)k$$

$$\text{New X Score} = [1200 + (1 - 0,76)] \times 100$$

$$\text{New X score} = 1224$$

$$\text{New Y Score} = \text{Previous Rank Score} - (\text{Score added by the winner})$$



$$\text{New } Y \text{ Score} = 1000 - 24 = 976$$

Source: The author

In this example,  $p$  is the win probability and  $k$  is a constant, which will be discussed later. The IAI ranking companies would start at the same score of 1500 points, since they will have no history of past disputes at  $t_1$ , and then carrying their history overtime.

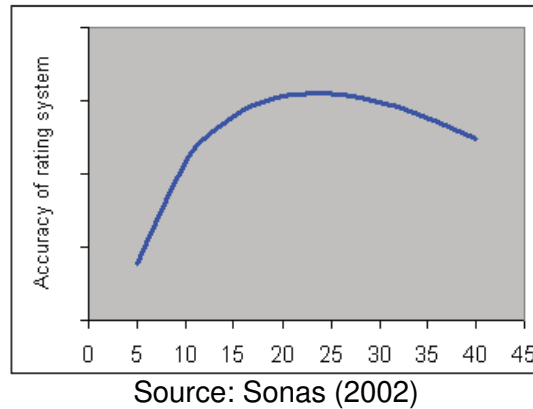
Table 1 - Difference in Elo-rating and the corresponding win expectation

Rating difference	Expected chance of winning	Difference	Chance	Difference	Chance
0>=dif<=3	0.50	122>=dif<=129	0.67	279>=dif<=290	0.84
4>=dif<=10	0.51	130>=dif<=137	0.68	291>=dif<=302	0.85
11>=dif<=17	0.52	138>=dif<=145	0.69	303>=dif<=315	0.86
18>=dif<=25	0.53	146>=dif<=153	0.70	316>=dif<=328	0.87
26>=dif<=32	0.54	154>=dif<=162	0.71	329>=dif<=344	0.88
33>=dif<=39	0.55	163>=dif<=170	0.72	345>=dif<=357	0.89
40>=dif<=46	0.56	171>=dif<=179	0.73	358>=dif<=374	0.90
47>=dif<=53	0.57	180>=dif<=188	0.74	375>=dif<=391	0.91
54>=dif<=61	0.58	189>=dif<=197	0.75	392>=dif<=411	0.92
62>=dif<=68	0.59	198>=dif<=206	0.76	412>=dif<=432	0.93
69>=dif<=76	0.60	207>=dif<=215	0.77	433>=dif<=456	0.94
77>=dif<=83	0.61	216>=dif<=225	0.78	457>=dif<=484	0.95
84>=dif<=91	0.62	226>=dif<=235	0.79	485>=dif<=517	0.96
92>=dif<=98	0.63	236>=dif<=245	0.80	518>=dif<=559	0.97
99>=dif<=106	0.64	246>=dif<=256	0.81	560>=dif<=619	0.98
107>=dif<=113	0.65	257>=dif<=267	0.82	620>=dif<=735	0.99
114>=dif<=121	0.66	268>=dif<=278	0.83	dif>=736	1.00

Source: Albers and Vries (2001)

Sonas (2002) research shed light to the of  $k$ , in his study was observed 226,00 chess games during the years of 1994 and 2001, to verify if the constant could change the predicting accuracy of a future chess game. The conclusion reached was the new scores are too sensitive to recent facts and the  $k$  starts to diminish the model accuracy after a maximum value. The graph below shows the curve reached on the research.

Graph 1 - K-Factor used in Elo Rating



Some researchers stated that the true asymmetry value is known by the company, since they possessed all the information, also determined what will be disclosed to the market. However, the level of information asymmetry varies hugely among companies, even the ones who follow the same disclosure and corporate governance protocol. Thus, managers don't have full control of the information asymmetry value of their own companies. It might take place since they might not be awarded over the market interest and if the investors truly understand the information disclosed by the company.

Verrechia (1983) stated that there is an equilibrium of asymmetry information which is carefully decided by managers and the market. Managers have the incentive to withhold information, especially the negatives ones, and traders are aware of it, until a certain limit. The absence of information might lead investors to misprice the company's stocks, which is not the goal of any manager.

It is a possible situation that a company intends to be transparent, adopting the best practices of corporate governance, but instead of decreasing the information asymmetry, it increases it because investors don't perceive its transparency. Although it may be true a situation where investors are believing in a companies' transparency, despite the company not adopting a full disclosure policy to the market. This last situation might be rapidly corrected by the index.

Hence, the value of information asymmetry is too sensitive over the analyst's perceptions of company disclosure, not even only in the present days, but when it comes to future guidance too. Corporate communication goes from managers to intermediates, investors and savers. It can take place by different sources, directly through financial reports, press releases and media, or indirectly through financial intermediaries and

financial analysts. Despite this, one of the main roles of corporate disclosure is to eliminate agency's problems (Healy & Palepu, 2001). For this reason, this work aims to understand how the analysts build their perceptions and which proxies are the most relevant to capture the value of a company's information asymmetry.

### **3.2 Independent Variable**

Literature has brought so far, an extensive enhance to proxy information asymmetry, as it can be perceived over different formats and degrees impacting cost of equity, pricing, stock price volatility and others. In order to prove the IAI consistence this work will segregate proxies based on Clarke and Shastri (2001), divided in three groups: external analysis, internal analysis and market microstructure. The next subsections will exploit those proxies.

#### **3.1.1 External Analysis Proxies**

An analyst of financial market uses information provided by the company to make forecast about a firm's future reports. This information is used to overcome financial reports (quarterly and annual), investors relation events and other forms of firm's communication. In addition, a great analyst would study the company's industry and its competitors. Hence, analysts' perceptions and recommendations (buy, hold or sell stocks) are an important source of information for an investor. Healy and Palepu (2001) found indicatives that analysts forecast, and recommendations add value to capital market, companies with greater coverage rapidly adjust their stock price due to new information. Although there are evidences that an analyst forecast can affect the stock price if they are biased. In the Brazilian market, the selling side usually issues companies' figures individually, stock price recommendations and future results.

This work will follow Shawn (2002) in our external analysis proxies for two reasons. The analyst forecast is verified the month before the actual earnings release, by proxying in this short term the biased optimisms are avoided, consistent with Brown et al. (1985). In addition, errors in forecast made very close to earnings announcement are associated with a firms-specific information rather than economy, or an industry's miss information.

The variable **COVERAGE** represents the number of sell side analysts on the Bloomberg data base covering the company. Lang and Lundholm (1993) found that companies with the best disclosure practices have larger analyst following, as well as less analyst forecast dispersion and less volatility. Consequently, it is expected that the higher the number of analysts following the firm, the lower would be the information asymmetry, and as lower would be the IAI score. As **COVERAGE** is a discrete variable, it could bias the model. Also, it hasn't a great variance over time (specially in short term, of course there is some companies which analysts drop coverage due to bankruptcy, merger or acquisition, or even because investors interest diminished). Hence, this work will adapt this proxy by calculating the distance among mean (of all sample) and standard deviation (of the company from the general mean). In this sense is possible the capture if a company attracts more or less sell side interest in the Brazilian context.

**ERROR** which simply is the absolute difference between the actual earnings and the Bloomberg median forecast. As literature suggested, higher differences are attributed to companies with higher degree of information asymmetry, hence it is expected that those companies appear with low score in the IAI.

Dierkens (1991) studied the importance of information asymmetry for firms during the process of equity issuance. The paper defined information asymmetry as a determination by assets' characteristic and manager and market behavioral. By proxy information asymmetry surround equity issue, she used the standard deviation of the daily stock price abnormal return for the subsequent year of issuance, the ratio of numbers of outstanding shares traded before and after the issuance, a dummy for public announcements and for growth opportunities proxy the ratio of market value of the equity and the book value of the equity.

Adapting Dierkens (1991) proxies, **VOL** will be used as a proxy for asymmetry information measured by the standard deviation of daily stock price variation of firm *i*. This proxy can be associated with the number of uninformed traders presented in the firm. As suggested by Wang (1993), the greater the percentage of informed traders, the greater the stock price volatility will be. It is expected that companies with higher levels of volatility will have higher levels of information asymmetry, and also lower scores on IAI ranking.

Table 2 - External Analysis Proxies

Proxy	Formula	Expected Result	Reference
COV	$\sigma_n/\mu$	(+)	Lang and Lundholm (1993)
ERROR	$AE - mdFore$	(-)	Shawn (2002)
VOL	$\sigma_{Stock.Price}$	(-)	Dierkens (1991)

Source: The Author

### 3.1.2 Internal Analysis Proxy

An extensive group of researchers dedicated to study information asymmetry on companies through their activities in the capital market, i.e. stock issue, debt issue. Its moment is particularly important because firms engage in roadshow and investor conference to increase voluntary disclosure and private channel communication, targeting analysts and investors with publicly available presentation (Schiemann et al., 2010). Focusing in amplifying transparency, companies aim to decrease opacity, hence decreasing cost of capital, bid-ask spreads and increasing market liquidity (Diamond and Verrecchia, 1991).

The bid-ask spread (BaA) was used as a measure of information asymmetry in most research projects, namely Chung (2006), Kanagaretnam, Lobo, and Whalen (2007), Chen, Chung, Lee, and Liao (2007), Wang and Zhang (2009), Chu and Song (2010), and Fauver and Naranjo (2010). The rationale of using the bid-ask spread can be obtained from Glosten and Milgrom (1985), who consider that argument spreads are consequences of asymmetric information among market participants. On the other hand, Huang and Stoll (1997) find that the bid-ask spread can be broken down into the cost of processing orders, carrying costs, and the cost of adverse selection. However, according to the authors, the most important part in determining the bid-ask spread is the cost of processing. Moreover, the intuition of using the bid-ask spread as a proxy to measure the asymmetry of information comes from the concept of Diamond and Verrecchia (1991), in which asymmetric information reduces the liquidity of the share. The bid-ask spread can be used as a measure of liquidity of an action, and it would also be a measure of information asymmetry; however, the fact that information asymmetry decreases the liquidity of a share is not the only factor that impacts liquidity and, consequently, the bid-ask spread.

Also, GO will be a measure of growth opportunities given by ratio of market value to book value of equity, as suggested by Smith and Watts (1992), McLaughlin et al. (1998). Besides, leverage has impacts on market-to-book measures, Penman (1996)

argue that the market can interpret higher level of leverage as a risk factor which has impacts in the market value. Literature suggests other problems with this proxy, the accounting data is quarterly bases, and higher levels of market-to-book can be associated with monopoly power, not growth opportunities. Besides the different base among market and accountability information, it still can measure the presence of opacity attributed to the discounted required by the investor to acquires firms' stock. This work particularly disagrees with this last concern. Besides, a monopoly has obvious advantages to the company in terms of market value, if investors are willing to pay a premium to its advantage, the price would be higher, decreasing the expected return. Hence, investors are willing to buy stocks at a higher price than accounting, if their perception is that the company will grow in the following quarters, enhancing its value. It is expected a higher level of market-to-book ratio associated with lower level of information asymmetry and higher IAI score.

Table 3 - Internal Analysis Proxy

Proxy	Formula	Expected Result	Reference
GO	$\frac{\text{market value of equity}}{\text{book value of equity}}$	(+)	Smith and Watts (1992)
BaA	$\frac{\text{Ask price} - \text{Bid price}}{\text{Price}}$	(-)	Diamond and Verrecchia (1991)

Source: The Author

### 3.1.3 Market Microstructure Proxy

The presence of information asymmetry is directly correlated with the presence of private information on the market. If there are investors more informed than others, the prices and stock return will be critically determined by its presence. By modelling stock liquidity and the frequency of bid and ask spread, it is plausible to make inferences about the likelihood of informed trading.

Easley and O'Hara (1987) developed a model which consists on measuring the impact in terms of price and size order of information asymmetry. Although they already said that the price-trade size relationship isn't determined exclusively by information effects, it is impacted by asymmetry issue. Their model identifies that in certain market conditions, the informed traders would trade solely large trade sizes, hence small traders are incapable of determining the asset price. They also noticed that market makers do

not know if they are dealing with informed, or uninformed trader, furthermore, there is always an uncertainty when it comes to the possibility of the existence of new information. As a result, there is a partial price equilibrium in which a large number of informed traders have small effects on the price. They also noted that informed traders maximize the expected profit of each trade individually, not in aggregate terms, what the authors called the competitive behavior.

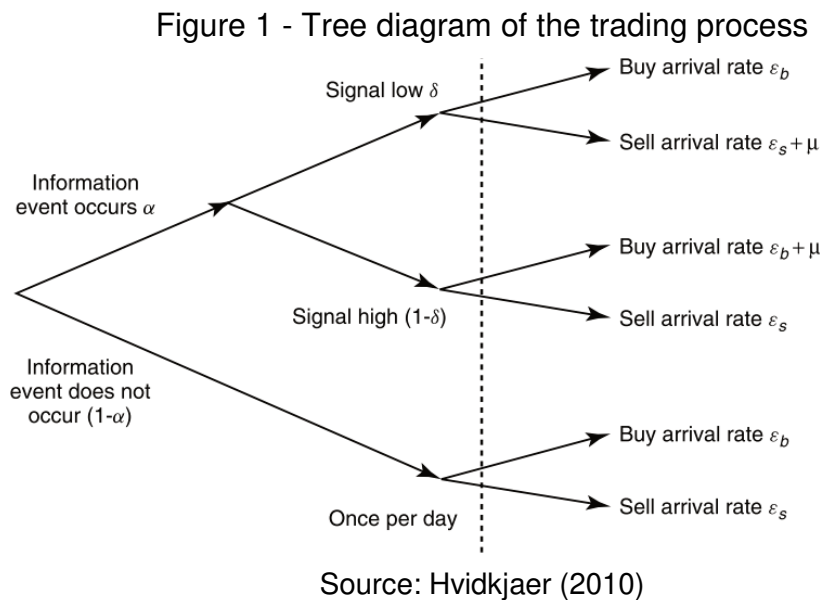
The existence of uninformed investors is the roots of information asymmetry and a plausible reason to market imperfection. Wang (1993) exploited a dynamic model of asset pricing under asymmetric information, identifying that uninformed traders contributes with market volatility. Investors are concerned about future cash flow, and noise traders, to determine stock price, when investors are less informed about company expectation of dividends growth rate, it becomes a harder task. In order to diminished noise traders, investors demand higher premium, turning prices more elastic to supply shocks. There is a positive relation between the existence of uninformed traders and higher premium, as information became less spread, stock price will not reflect companies' fundamentals, increasing risk premium.

Because literature shows that traders demand higher return to invest in companies which have greater private information, Easley, Kiefer and O'Hara (1997), Easley, Hvidkjaer, and O'Hara (2002) (EHO), Easley and O'Hara (2004), develop an equilibrium model to price information asymmetry, shedding light in the assumption that well informed investors play a critical component on the price formation process (Hvidkjaer, 2010). Since informed investors are capable to rebalance their portfolios when news information arrives, uninformed investors are always on the wrong side, holding too many stocks with bad news and a few with good news, in a frustrated attempt to diminish the risk by diversification. Hence, the presence of private information in the market shows that the CAPM theory is wrong about systematic risk, investors won't hold identical portfolios because the expected return and risk perception aren't the same among them, uninformed investors require a greater return to hold stock and are more sensitive to new information. On the other hand, because informed investors know which stocks have good and bad news, they are able to hold (or even sell bad stocks) for a longer time.

As a proxy for information asymmetry, Easley, Engle, O'Hara y Wu (2008) continue the studies and develop a dynamic model measuring the interaction among informed and uninformed investor in terms of liquidity, market depth and order flow

through time. Literature uses the probability of informed trading (PIN) as one of the most accurate proxies to measure asymmetry. It is based on the theoretical assumption that informed traders are the ones who unbalance the trade equilibrium, and by using their private information it's possible to infer that abnormal return is plausible, since they are always on the right side of the trade, causing the adverse selection problem to uninformed investors.

This work will follow the contributions of Ealey and O'Hara and Hvidkjaer model to determine the probability of informed trading. Hvidkjaer (2010) suggested a tree diagram of trading process which good news  $\delta$  or bad news  $(1 - \delta)$  occur with  $\alpha$  probability at a date  $t$ , changing the stock price to  $\bar{V}_t$  if there is good news arriving, or to  $V_t$  if it is bad news, as suggested in the figure below.



During a trading day, investors place their orders according to a Poisson process executing them according to their own necessity. Informed investors arrive at rate  $\mu$  as uninformed investors –  $\varepsilon_b$  for buyers and  $\varepsilon_s$  for sellers – trade for liquidity reasons. Estimating via maximum likelihood is possible to determine the PIN of stock  $j$  at date  $t$ . With the help of Bloomberg platform, the number of buyers and sellers of a day gives the first step of the estimation. The following equation is the likelihood formed by these investors, where  $B$  is the total number of buyers,  $S$  is the total number of sellers and  $\theta = (\mu, \varepsilon_b, \varepsilon_s, \alpha, \delta)$  is the parameter vector. As suggested by the diagram, this likelihood function is weighted by the probability of good news taking place, bad news taking place, or even no news at the date, hence this likelihood function is a mixture of three Poisson



probabilities, weighted by the probability of having good news  $\alpha(1 - \delta)$ , bad news  $\alpha\delta$ , and no news  $(1 - \alpha)$ .

Equation 2: The likelihood for the total number of buys and sells

$$\begin{aligned} \mathcal{L}\left((B, S) \mid \theta\right) &= \alpha(1 - \delta)e^{-(\mu + \varepsilon_b + \varepsilon_s)} \frac{(\mu + \varepsilon_b)^B (\varepsilon_s)^S}{B! S!} \\ &+ \alpha\delta e^{-(\mu + \varepsilon_b + \varepsilon_s)} \frac{(\mu + \varepsilon_b)^B (\varepsilon_s)^S}{B! S!} \\ &+ (1 - \alpha)e^{-(\varepsilon_b + \varepsilon_s)} \frac{(\varepsilon_b)^B (\varepsilon_s)^S}{B! S!} \end{aligned}$$

Source: The Author

Following Hvidkjaer (2010) in order to increase computing efficiency and to reduce truncation error, the likelihood function as rearranged to the following equation, where  $M_t = \min(B_t, S_t) + \max(B_t, S_t)/2$ ,  $x_s = \varepsilon_s/\mu + \varepsilon_s$ ,  $x_b = \varepsilon_b/\mu + \varepsilon_b$ ,

Equation 3: The log likelihood for the total number of buys and sells

$$\begin{aligned} \mathcal{L}\left((B_t, S_t)_{t=1}^T \mid \theta\right) &= \sum_{t=1}^T \left[ -\varepsilon_b - \varepsilon_s + M_t (\ln x_s) + B_t \ln(\mu + \varepsilon_b) + S_t \ln(\mu + \varepsilon_s) \right] \\ &+ \sum_{t=1}^T \ln \left[ \alpha(1 - \delta) e^{-\mu} x_s^{S_t - M_t} x_b^{-M_t} + \alpha\delta e^{-\mu} x_s^{-M_t} + (1 - \alpha) x_s^{S_t - M_t} x_b^{B_t - M_t} \right] \end{aligned}$$

Source: The Author

After the estimating parameters  $\theta = (\mu, \varepsilon_b, \varepsilon_s, \alpha, \delta)$ , EHO model suggested that it is time move to its economic usage. As suggested by the literature, market makers would set the price expecting losses when trading with informed traders and gains when trading with uninformed traders. This result between gains and losses is the source of the bid and ask spread, expressed in the following equation, where  $\alpha\mu + \varepsilon_b + \varepsilon_s$  is all arrival rates for all orders (informed and uninformed traders) and  $\alpha\mu$  is the arrival rate

for informed traders only. As a result, PIN will demonstrate if the stock congregates higher or less proportions of informed traders.

Equation 4: PIN

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s}$$

Source: The Author

The PIN estimation can over post the conceptual of bid and ask spreads as they came from the same theory, nevertheless, it is more robust Easley et al (2002). In addition, PIN has an extensive impact on companies: higher cost of capital Duarte, J., Han, X., Harford, J. & Young, L. (2008), presence of insider trading Aslan, H., Easley, D., Hvidkjaer, S. & O'Hara, M. (2011) and higher expected returns Easley, D. & O'Hara, M. (2004). It is suggested that as higher PIN higher is higher would be information asymmetry. Hence, it is expected that higher levels of PIN would be associated with lower score on IAI ranking.

Table 4 - Market Microstructure Proxy

Proxy	Formula	Expected Result	Reference
PIN	$\frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s}$	(-)	Easley, Hvidkjaer, and O'Hara (2010)

Source: The Author

### 3.3 Control Variable

Free float and  $\ln(\text{tradedvalue})$  was included as control variables, as also presented in the literature as proxy of information asymmetry.

Free Float is a proxy for corporate control. The presence of insiders, e.g. control blocks, usually have greater access to private information (Leuz & Verrechia, 2000), regarding that free float increases public information, due to more investors having access to private information. Hence, free float is associated to lower levels of information asymmetry and higher levels of IAI. The descriptive statistic shows that

Brazilian companies, which belong to IBrA, have 60% means of free float, with a high standard deviation of 25%.

Traded Value is the total amount traded in the security's currency. This value represents all traded prices, multiplied by the number of shares relating to each price. It proxies the investor's demand of the company, naturally whether traded value increases, it is due to demand's increase. As seen in the literature, investors traded given new information arising, or by liquidity reasons. Hence more information, which naturally decrease information asymmetry, increasing traded value. It is expected that the higher the traded value, the higher IAI would be.

Table 5 - Control Variable Proxies

Proxy	Formula	Expected Result	Reference
Free Float	%ff	(+)	(Leuz & Verrechia, 2000)
Traded Value	$\ln(\text{tradedvalue})$	(+)	(Huddart, 2007)

Source: The Author

### 3.4 Econometric Model

The purpose of this research is to create a model with secondary data, presented above divided into three groups: Internal Analysis, External Analysis, and Market Microstructure. The dependent model is the Information Asymmetry Index which reflects the analysts' perception over companies' disclosure.

IAI was conducted four times since from 2016-2018. Because the aim is to analyze data cross sectionally and over time, since information asymmetry varies over companies and into a same company over time, a panel data is the desired model. Despite the theoretical assumption, the data doesn't have a long observation in terms of time, neglecting it could lead to wrong bias, for this reason the model choice is a pooled cross section with dummies for time and for sector.

#### Equation 5

$$IAI_j = \alpha + \beta_1 COV_j + \beta_2 ERROR_j + \beta_3 VOL_j + \beta_4 GO_j + \beta_5 BaA_j + \beta_6 PIN_j + \beta_7 D_i + \beta_8 D_j + \varepsilon$$

where:

*COV* = Proxy for Coverage, quantity of sell side analyst

*Error* = Proxy for the absolute difference between EPS median forecast and EPS realized

*Vol* = Volatility diary basis for stock

*GO* = Proxy for Growth Opportunities, it is the market-to-book ratio

*PIN* = Probability of informed Trader

$D_i$  = Dummy for time

$D_j$  = Dummy for sector

The Variance Inflation Factor (VIF) was conducted in order to test if multicollinearity is presented on data. The VIF is a factor which represents the relation among the variable  $\hat{\beta}_j$  and other explanatory variables, if VIF is high, it increases the variance of the estimators and indicates the presence of collinearity. Literature suggests that if the VIF of a variable exceeds 10, it is highly collinear. Another measure is TOL, which is the opposite of VIF, as closer to zero the measure is, the greater the collinearity. The test does not reflect the presence of multicollinearity in the model, appendix shows the test results

The crossed product of White test was conducted in order to avoid Heteroskedasticity and specification bias, despite that it consumes many degrees of freedom when a model has many estimators. The test results a chi-squared value less than the critical chi-square at 5% of significance, can't reject the null hypothesis of Homoscedasticity, the appendix shows the test result.

Regarding the presence of outliers, the independent variable will be minorized on the minimum level at 2.5% and on the maximum level at 97.5%, using Stata resources. Also, will be testes the leverage contamination trough robust regression and Cook's distance.

## 4 OUTCOME ANALYSIS

This chapter begins with the presentation of the descriptive statistic and correlation matrix of the data, which are composed by 120 held public companies during the period of 2016 to 2018. Additionally, it presents the outcomes of the econometric model and the executed tests.

### 4.1 Statistic Descriptive

The starting point is to collect the analyst answers to build the IAI. An e-mail was sent to many asset managements, investment banks and stock brokers, asking them to answer the following question: “Which of the following companies have the best disclosure level?”, as well as if there were any CFA or CNPI holders. Before answering the ten disputes pairs, a brief explanation of the meaning of disclosure was given which follows: “Communication of all information, positive and negative, about a company. The aim is to provide information to debtholders and shareholders, letting them have an opinion about the companies’ financial situation. Appendix provides a screenshot of the website. The 120 companies that participated in the survey belong to Brazil Broad-Based Index, besides “Ibovespa” Index is the most usual in the Brazilian market, it is composed only by the 65 companies with higher liquidity, which could bias the sample, since the aim of this work is to provide a broad index of Asymmetry Information.

Table 6 shows the descriptive statistics of the IAI by collection and aggregation. The standard deviation is increasing overtime, which is expected, since companies carry past disputes, hence the range points are widening overtime, the mean is around 1500 points, which makes sense given the index methodology, i.e. in a direct dispute, a company loses points to a winner.

**Table 6 - Descriptive Statistic IAI.** The table presents the descriptive statistic for the Information Asymmetry Index for all period individually and aggregated. Columns show the number of observations, mean, standard deviation, median, minimum, maximum and the number of respondents.

IAI	Obs	Mean	Std. Dev.	Median	Min	Max	Respondents
$t_1$	104	1503	50.43	1508	1280	1633	41
$t_2$	107	1508	70.97	1507	1355	1716	52
$t_3$	120	1506	82.41	1501	1341	1718	37

$t_4$	115	1502	88.88	1495	1218	1731	34
Total	446	1505	75.06	1502	1218	1731	164

Source: The Author

Almost every secondary data came from the Bloomberg except PIN which is provided directly from B3, the Brazilian stock market, with the auxiliary of R software. To estimated PIN two R packages were used, first Perlin and Ramos (2006) developed the GetHFDData which import and aggregate high frequency trading data straight from B3' ftp website after, the package developed by Celik and Tiniç (2017) was used to infer the probability of informed trader based on EHO model and Lin and Ke (2011) (LK factorization). Unfortunately, B3 is limiting the access through ftp web site, formerly data was available for a longer period, however due to changes on internal policies, the stock market is limiting the access just for the last six months. The author downloaded data from July 2<sup>nd</sup> 2018 to December 28<sup>th</sup> 2018 and infer PIN for this period, presented latter. Because there isn't data available for the role period surveyed, Market Microstructure proxy was substituted to Bid and Ask spread, which is not exactly the same, but congregated some similarities: are independent form regulated or voluntary disclosure, has no researcher bias and higher levels of PIN and Bid and Ask Spread are associated with higher levels of information asymmetry.

Table 7 details the information asymmetry proxies. There is some missing value, especially from the Error proxy. It comes from Bloomberg median EPS, which is formed by the sell side analysts that published their forecast on the software. Besides coverage has a higher number, not all of them published the analysis. Because Coverage is a discrete variable, which could disrupt predictive model capability, it was modified by an index resulting by the division among the mean (per period) and the standard deviation (per period and per company), hence the variable is measuring if the company  $j$  receives more, or less interest from the sell side, on the average from time  $i$ . The proxy BaA refers to Bid and Ask spread, it is the average of all bid and ask spreads taken as percentage of the mid-price.

**Table 7 - Descriptive Statistic.** The table presents the descriptive statistic of all the proxies and the independent variable. Columns show the number of observations, mean, standard deviation, median, minimum and maximum.

Proxies	Obs	Mean	Std. Dev.	Median	Min	Max
Cov	443	0.00	0.08	0.01	-0.17	0.13
Error	197	-0.19	0.67	-0.12	-1.72	1.07
Vol	445	30.93	9.68	29.27	16.88	52.87
Go	431	2.67	2.36	1.73	0.54	9.16
BaA	446	0.17	0.11	0.13	0.06	0.53
ff	446	60.90	25.08	54.25	27.55	100.00
Intradedvalue	446	16.36	1.66	16.53	13.24	19.08

Source: The Author

## 4.2 Correlation Matrix

Table 8 shows the correlation matrix of the independent variable. Because they are all proxy's information asymmetry, it was expected a high correlation, especially on those proxies from the same group, e.g.: Coverage and Error. Besides that, correlation wasn't high with few exceptions, Coverage with  $\ln(\text{tradedvalue})$  is positive correlated at 0.34 ( $p < 0.05$ ) showing that Coverage increase investors' demand, consistent with the literature. This example is agreed with finance literature, bigger companies have more coverage which increase the traded volume either is plausible to assume that because companies are big, traded value would be high too, which attract attention of sell side. Independent of the discussion of cause and consequence, for this work is sufficient the concept that those are good information asymmetry proxies. Also, Bid and Ask spread presents a negative correlation with  $\ln(\text{tradedvalue})$  at -0.38 ( $p < 0.05$ ) and Coverage presents negative correlation with Bid and Ask spread at -0.48 ( $p < 0.05$ ). Again, those examples follow literature, companies with higher coverage tends to have less information asymmetry and consequently fewer Bid and Ask Spread. The same as traded value, com with bigger values traded tends to have less information asymmetry and lower Bid and Ask spread.

**Table 8 - Matrix Correlation.** This table presents the correlation among all the independent variable. Note \*, \*\*, \*\*\* represents 5%, 1% and 0,1% statically significance.

	Cov	Error	Vol	Go	BaA	ff	Intradedvalue
Cov	1						
Error	0.0448	1					
Vol	0.1251*	0.0797	1				
Go	0.1817*	0.081	-0.1393*	1			
BaA	-0.4880*	0.0744	0.2686*	-0.1848*	1		
ff	0.056	-0.0468	-0.0156	-0.1376*	-0.1929*	1	

Intradedvalue	0.3412*	-0.2220*	-0.1166*	0.1023*	-0.3846*	0.1922*	1
---------------	---------	----------	----------	---------	----------	---------	---

Source: The Author

### 4.3 Outcomes Presentation

The model aim is to outcome a high R-squared ( $R^2$ ) to verify that the Index is capturing the sensitivities of market analysts with respect to their perceptions about the informational asymmetry of Brazilian companies.

Initially, a scatter plot was made in order to verify if the independent variables were correlated in agreement with the theory. Coverage, Go, Free Float and Traded Value are related to higher points on IAI ranking, while Error, Volatility, Bid and Ask Spread and PIN contributes to lower points on IAI ranking, as predicted on the literature. The graph is at Appendix.

The pooled cross-sectional model, presented on table 4, was tested by multicollinearity, heteroskedasticity and for outliers and leverage contamination. VIF test was conducted to test over multicollinearity, it outcomes no collinearity among variable, literature suggests that low collinearity must have factor under 10, while the variable means is 1,76. The Breusch-Pagan and White test for Heteroskedasticity results, by not being possible to reject the null hypothesis, that the error variance is constant, i.e. the data is homoscedasticity. For outliers and leverage contamination, the robust regression and Cook's distance were tested. Besides, some companies presented high leverage (Cielo, CPFL Energia, ItaúUnibanco, PetroRio, JBS and Comgas) weighted and reweighted least squares regression don't show large absolute residuals, hence it is not necessary to run robust regression.

**Table 9 - Polled Cross Section – All Sample 2016-2018.** Model (1) presents all independent variables with control variables for all the period. Model (2) is similar previous model except for variable Error, which were dropped. Model (3) includes dummies for time. Model (4) presents dummies for sector. Note \*, \*\*, \*\*\* represents 5%, 1% and 0,1% statically significance.

VARIABLES	(1) IAI	(2) IAI
Cov	179.29*** (2.63)	144.39*** (3.11)
Error	-8.11 (-1.01)	
Vol	-1.91*** (-3.51)	-1.24*** (-3.53)
Go	7.17***	6.26***



	(2.91)	(4.36)
BaA	-103.60*	-81.52**
	(-1.70)	(-2.22)
ff	0.47**	0.23
	(2.23)	(1.64)
Intradedvalue	4.67	5.14**
	(1.50)	(2.35)
Constant	1,467.20***	1,442.09***
	(26.20)	(36.39)
Observations	186	428
R-squared	0.3108	0.2245
Adj R-squared	0.2837	0.2134
Period	2016-2018	2016-2018

Source: The Author

Table 9 present two regressions. First presents all independent variables, the signals are as predict by literature and by the graph presented. *Cov* (Coverage Index) is significant ( $p < 0.001$ ) with high coefficient, although *Error* (Difference between Earning per share realized by the companies and the median estimative of sell side analysts provided by Bloomberg) wasn't significant. *Vol* (Volatility) is significant ( $p < 0.001$ ) with low coefficient and *Go* (Market-to-book ratio) is significant ( $p < 0.001$ ) with a small coefficient too. *BaA* (Bid and Ask spread) isn't significant ( $p < 0.05$ ) with high coefficient. The control variables: *ff* (free float) are significant ( $p < 0.01$ ), again, with a very low coefficient;  $\ln(\textit{tradedvalue})$  isn't statically significant. The adjusted R-squared 0,28, which is good for cross-sectional models. The number of observations is fewer because *Error* measure has missing value. Besides, the model presents few variables with no significance. *Cov* is the grand variable of the model, it is the one which better explains how companies can have higher levels on IAI, and consequently lower levels of information asymmetry. Also, *BaA* has a high coefficient to explain IAI score, although because its level of statistical significance, its value is damaged.

In the model 2, the variable *Error* was omitted to enhance the number observations (to 186 for 428). The difference of observations is due to Bloomberg platform and doesn't have analysts forecast for many companies, even if they have coverage. *Cov* is significant ( $p < 0.001$ ) with high coefficient, comparing to model 1, it remains statistical significance, but with lower coefficient. *Vol* presents a similar result in the model 1, significant ( $p < 0.001$ ) low coefficient as the same as *Go* which is significant ( $p < 0.001$ ) with small coefficient. *BaA* proxy improved its statistical significance ( $p < 0.01$ ) from the last model, although the coefficient decreases. The

control variables: *ff* are not significant;  $\ln(\text{tradedvalue})$  is significant ( $p < 0.01$ ) with a small coefficient. The adjusted R-squared decrease from model 1, besides an independent variable was excluded, which could enhance R-squared, in fact apparently *Error* is an important variable to the model, for this reason in the following model it will not be excluded.

**Table 10- Polled Cross Section – All Sample 2016-2018.** Model (3) includes dummies for time. Model (4) presents dummies for sector, and model (5) includes all the dummies for sector and for time. Note \*, \*\*, \*\*\* represents 5%, 1% and 0,1% statically significance.

VARIABLES	(3)	(4)	(5)
	IAI	IAI	IAI
Cov	140.40** (1.98)	191.83*** (2.73)	159.77** (2.16)
Error	-11.63 (-1.44)	-10.06 (-1.30)	-12.60 (-1.62)
Vol	-2.29*** (-4.07)	-1.79*** (-3.15)	-2.06*** (-3.52)
Go	6.72*** (2.73)	9.22*** (3.55)	8.65*** (3.35)
BaA	-41.20 (-0.60)	-44.60 (-0.72)	-2.37 (-0.03)
ff	0.45** (2.13)	0.49** (2.23)	0.48** (2.19)
Intradedvalue	15.81** (2.60)	2.41 (0.79)	10.96* (1.80)
t2	27.80 (1.45)		31.15 (1.67)
t3	33.58 (1.63)		35.58 (1.80)
t4	70.56* (2.59)		61.48* (2.32)
sec1		30.92 (0.81)	26.42 (0.69)
sec2		26.15 (1.06)	25.21 (1.02)
sec3		-8.48 (-0.42)	-6.96 (-0.34)
sec5		46.13* (2.35)	46.46* (2.35)
sec6		67.68** (3.02)	64.10** (2.83)

sec7		-36.98	-40.68
		(-1.32)	(-1.43)
sec8		-2.05	-1.95
		(-0.05)	(-0.05)
sec9		7.94	13.41
		(0.19)	(0.32)
sec10		15.98	23.10
		(0.50)	(0.73)
sec11		11.62	12.66
		(0.52)	(0.57)
Constant	1,248.06***	1,465.47***	1,289.24***
	(11.35)	(25.58)	(11.80)
Observations	186	186	186
R-squared	0.3371	0.4158	0.4360
Adj R-squared	0.2992	0.3567	0.3676
Period	2016-2018	2016-2018	2016-2018

In the model 3, it was included dummies representing the periods when the survey took place (October/16, March/17, December/17 and November/18). *Cov* is significant ( $p < 0.01$ ), which decreased since the past two, and the coefficient remains high; *Error* isn't significant. *Vol* is significant ( $p < 0.001$ ) with a small coefficient, similarly with the previous one, *Go* is statically significant ( $p < 0.001$ ) with a small coefficient, again similarly with the previous one. *BaA* is not significant with this group of proxies. The control variables are both significant, *ff* ( $p < 0.01$ ) with a low coefficient and  $\ln(\text{tradedvalue})$  is also significant ( $p < 0.01$ ) with a representative coefficient. Was included dummies ( $t_2, t_3, t_4$ ) for three periods (March/17, December/17 and November/18) aiming to identify whether time contributes to enhancing IAI companies' score. Although they all seem to increase IAI companies' level, just dummy  $t_4$  is significant ( $p < 0.05$ ) with a representative coefficient, i.e. companies tend to have a higher IAI level over time. Notice that R-squared increases with dummies for time.

In the model 4, dummies were included, representing the eleven sectors, the classification is made by B3. *Cov* again, it is significant ( $p < 0.001$ ) with high coefficient, *Error* isn't significant just as model three, *Vol* is close to last model, significant ( $p < 0.001$ ) with low and *Go* significant ( $p < 0.001$ ), with higher coefficient since the last model. *BaA* once again is not significant. For the control variables, *ff* is significant ( $p < 0.01$ ) with a very low coefficient,  $\ln(\text{tradedvalue})$  isn't significant. Dummies outcomes don't have much statistical significance, only sectors 5 and 6 are significant with relevant

coefficient. Getting into results, Finance and Materials sectors have more coverage than mean, higher IAI level, more free-float, bigger asset and smaller bid and ask spread, as shown on table 10. This characteristic would decrease information asymmetry over the company according to literature. Additionally, the finance sector is more regulated than the other, since besides CVM regulation, banks and financial companies are regulated by the Brazilian Central Bank, which could increase company's disclosure. As sector 4 is the largest sector, it was chosen to not have a dummy. It was also noticed that dummies sectors improve the R-squared significantly.

In the model 5, all the dummies for time and for sector were included, presented on models 3 and 4. *Cov* is significant ( $p < 0.01$ ) with high coefficient, as in the other models *Error* is not significant. *Vol*, as in the other models, is significant ( $p < 0.001$ ), but with low coefficient and *Go* is also significant ( $p < 0.001$ ) with a similar coefficient on the other models. The proxy *BaA* isn't significant. The control variables *ff* is significant with low coefficient ( $p < 0.01$ ) and  $\ln(\text{tradedvalue})$  is significant ( $p < 0.05$ ). The presence of all dummies made the 3 times dummies emphasize that it is an important variable to make better known companies and thus diminished their opacity. The sector dummies had very similar outcomes, in comparison to the last model, Material (sector 5) significant ( $p < 0.05$ ), also Financial (sector 6) is significant ( $p < 0.01$ ). Equally, on model 4, dummies increase the R-squared, the model has the highest factor comparing to previous models.

#### 4.4 Robustness Test

In order to verify the model efficacy, as well as the independent variable used, this subsection proposes the inclusion of two different variables: ADR issued as a dummy variable, which companies have had issued, should be best known on the foreign market, hence, given the predominance of foreign capital on Brazilian stock market, they may have lower levels of information asymmetry; and PIN, which literature shows as an one of the best proxies for information asymmetry – it might be present for all periods that the survey was conducted, instead of *BaA*, although B3 provides data just for the last six months. Table 11 and 12 present the outcomes.

**Table 11- Robustness test ADR inclusion.** In model (6), proxy for ADR issue was included. Model (7) contain the cross-variable ADRBaA. Model (8) has the cross-variable ADRCov. Note \*, \*\*, \*\*\* represents 5%, 1% and 0,1% statically significance.

VARIABLES	(6) IAI	(7) IAI	(8) IAI
Cov	187.51*** (2.71)	187.29*** (2.68)	187.17** (2.44)
Error	-7.96 (-0.99)	-7.93 (-0.98)	-7.95 (-0.98)
Vol	-1.94*** (-3.54)	-1.94*** (-3.52)	-1.94*** (-3.53)
Go	7.08** (2.86)	7.07** (2.84)	7.08** (2.85)
BaA	-105.07 (-1.72)	-104.39 (-1.62)	-105.08 (-1.72)
ff	0.47* (2.21)	0.47* (2.20)	0.47* (2.21)
Intradedvalue	4.93 (1.57)	4.91 (1.55)	4.92 (1.53)
ADR	-7.50 (-0.66)	-6.92 (-0.33)	-7.54 (-0.63)
ADRBaA		-4.52 (-0.03)	
ADRCov			1.56 (0.01)
Constant	1,467.32*** (26.16)	1,467.56*** (25.89)	1,467.44*** (25.61)
Observations	186	186	186
R-squared	0.3125	0.3125	0.3125
Adj R-squared	0.2814	0.2774	0.2774
Period	2016-2018	2016-2018	2016-2018

Source: The Author

The results are not significant for the ADR contribution to increase IAI level, not only in terms of statistical significance, but also in the sign of the variable, it was expected to be positive. The interest variable *Vol* is statistically significant with low coefficient, similarly previous model, while other variables aren't significant. In addition, given the importance of foreign investors on the Brazilian stock market, it was attempted to verify if ADR has effects on Coverage and on Bid and Ask spread from the interaction of cross-variables. However, again the outcomes are inconclusive due to statistical significance.

**Table 12 – Robustness Test PIN inclusion.** In model (9) PIN proxy is included and ran just for  $t_4$ . Model (10) presents the cross-variable PINcov, is also ran for  $t_4$ . Note \*, \*\*, \*\*\* represents 5%, 1% and 0,1% statically significance.

VARIABLES	(9) IAI	(10) IAI
Cov	-8.81 (-0.07)	579.71 (1.88)
Error	-10.21 (-0.74)	-10.84 (-0.80)

Vol	-2.75** (-2.87)	-2.68** (-2.85)
Go	8.89 (1.92)	9.47* (2.09)
ff	0.20 (0.56)	0.19 (0.53)
Intradedvalue	21.98* (2.25)	18.45 (1.90)
PIN	-425.20 (-1.49)	-494.23 (-1.76)
PINCov		-3,520.97* (-2.10)
Constant	1,313.16*** (8.99)	1,366.09*** (9.43)
Observations	81	81
R-squared	0.3051	0.3453
Adj R-squared	0.2385	0.2726
Period	2016-2018	2016-2018

Source: The Author

The results are inconclusive using PIN, instead of *BaA*, although the coefficient is negative, as predicted on the literature, the statistical significance prevents drawing conclusions from the model. Even *Cov* is not significant and with negative coefficient, contradicting the literature and previous models presented. Only *Vol* is significant ( $p < 0.01$ ). Because it was a negative surprise, given the robustness of PIN model, it was attempted to verify if the interaction of PIN and *Cov* are representative. The results with the cross-variable enhance the model, now *Vol*, *Go*, and *CovPIN* are significant. Again, Coverage contributes to improve IAI level and now the presence of informed traders on asset pricing decreases IAI score, as predicted on literature. It is interesting to notice how influent the *CovPIN* is to determine IAI. On the financial market, sell side analysts are paid to provide their opinion about the company, whether good or bad. This cross-variable demonstrates that the sell side opinion, combined with insider trading asset pricing decreases the companies' disclosure perceived by the investors.

## 5 DISCUSSION

Information Asymmetry is a wide research field in finance, involving many company's aspects regarding executive compensation, cost of capital, level of indebtedness, profitability, shareholder return, liquidity, control structure dividend policy, asset pricing and others. Many researchers have been proxying it over different forms, choosing the ones that best fit on their work. Based on the study of literature, this work adapted Clarke and Shastri (2001) proxy segregation, dividing those on three groups: external analysis, internal analysis and market microstructure.

### 5.1 External Analysis Outcomes

This group is highly important to determine IAI company's score, clearly the reason isn't the quantity of proxies, but because Volatility and Coverage has a high influence on IAI score, due to their coefficient and the mean value. Volatility has important impact due to coefficient size, Coverage is also important, but in a lower level.

Beginning with Coverage the results become better after the calibration proposed this proxy was adapted to capture the difference of each company and the average coverage, instead of using the quantity of analyst. In this sense, *Cov* measures the influence of being ahead the average on the IAI score. The finding shows that it is an important factor, grand positive (negative) difference from average are associated with higher (lower) score on IAI.

Lang and Lundholm (1993) already discusses over analyst coverage, although their conclusion has reversed the order established herein. For then, disclosure practices increase the demand for analyst reports. In fact, for this work it is relevant to notice that Coverage impacts on companies' asymmetry.

On the other hand, Li and You (2015) found that coverage creates value for companies but haven't found conclusive results for reducing information asymmetry. In time, ItaúUnibanco is the company with the highest IAI score (1731), but they aren't the company with mayor coverage, Banco do Brasil, for example, has 21 sell side analysts. Nevertheless, evidences here presented emphasize the importance of sell side analysts on company's asymmetry.

Given the importance of sell side analysts *Error* might be relevant as well, since the final product produced by the analysts are their companies' forecast. Also, literature

shows that proxy has the same impact of Coverage. Besides these assumptions, *Error* is not significant in any of the models presented. Under those circumstances, it is possible to infer that the variable has calibration problem: there are a lot of missing values on data base and the reason why is not clear, if Bloomberg doesn't feed correctly, or if analysts aren't willing to share their forecast (some stock brokers and investment banks are not allowed to share their forecast to everyone, just for clients). Although the lack of significance, in model 2 which ran without the variable, the R-squared was the worst of the models used, it shows that even though *Error* contributes for model efficacy.

Volatility was the unique proxy statistically significant in the models, consistent with Dierkens (1991), this proxy decreases company's IAI score. As discussed by Verrechia (1983), noise traders, or uniformed traders (Wang 1993) can mispricing a stock by the lack of information companies provide. As a company enhance its disclosure policy, more information market would possess, investors would demand a lower premium risk to buy the stock and volatility tends to diminish. The findings in this work agree with theoretical assumption, in the sense that higher IAI scores are associated with lower volatility. In time, the company with the lower score is PetroRio, also it has the highest volatility.

## 5.2 Internal Analysis Outcome

The proxy *Go* had an important contribution to this work, in almost all the model ran, it was statically significant. There are some theoretical discussions surrounding growth opportunities proxy. On the one hand, it might enhance information asymmetry since companies with grant investment, projects, or newer tends to expand their figures, enhancing profitability, hence combined with growth it is a possibility of failure. In this sense, investors might expect larger returns under this circumstance. In addition, authors criticized the proxy due to accounting data in quarterly bases, monopoly power and company leverage. Despite all those arguments, this work understands that market-to-book ratio is a measured that captures how much investors are willing to pay per stock in relation to accountability value. If investors buy stocks higher than book value, it means that they expect increasing returns, overcoming the risk taken. From this assumption and the statistical significance of the proxy, it is understood that when a



companies' market value is higher than the book value is because the disclosure provided made investors reduce uncertainties about future.

### 5.3 Market Microstructure Outcome

EHO model was expected to the grand variable of this work. Its methodology is ahead of others information asymmetry proxies, as the independent character from accountability rules, regulation disclosure, researchers' interpretation and bias, however, demands from the researcher greater econometric knowledge and access to data. When this researcher began, it wasn't expected not having this data, Bloomberg, which is one of the biggest market platforms has it, but it doesn't keep record. Another choice was to download straight from B3 via FTP. Again, for the researcher surprise, B3 limited access to only six previous months. Therefore, the data was available at the period of the last survey. Hence this proxy was extracted from the model and used just as robustness.

As  $G_0$ , PIN doesn't play an important part to determine IAI score due to its lack of statistical significance. Although, theoretically the variable sign was as predicted. Greater asymmetry information is associated with a larger bid and ask spread, on average, hence a lower IAI score.

Table 13- PIN Descriptive Statistic

Proxy	Obs	Mean	Std. Dev.	Median	Min	Max
PIN	115	0.1546	0.0444	0.1449	0.0953	0.4512

Source: The Author

When analyzing PIN starts comparing it with other Brazilian studies with this methodology. Martins and Paulo (2014) studied 194 companies over 2010/2011 finding that, on average, PIN was 25%, Bopp (2003) studied ADR from Brazilian companies on 2001 finding a probability of informed trader of 23,9%. This work uses data from July/2018 to December/2018 for 115 companies finding, on average, PIN of 15,45% (minimum of 9,53% and maximum of 45,11%).

PIN itself wasn't significant for the studied period, possibly if a longer time was available the outcomes might have been different. It is interesting to notice the significance of cross-variable PIN and Coverage, showing that sell side can increase

significantly information asymmetry if they perceived the presence of informed traders on companies' asset pricing.

#### 5.4 Dummies Outcome

As this work uses pooled cross-section, dummies for time were included in order to verify if the variable time affects IAI score. On models three and five, only  $t_4$  was significant, i.e. on average recent time increases IAI score. At this finding, it is plausible to assume that, on average investors aren't worried about past events, but are worried about what is happening at the present time. It's still possible to assume that investors demand time to know the companies well and reduces its opacity according with Boulton, Smart and Zutter (2011) which showed that IPO companies have a higher degree of opacity.

Equally important, this outcome might be attributed to the recent IAI history. As on the first collection all the companies start at score 1500 and then carry the disputes results over time. Index maturity might be reached at  $t_4$ .

Sector dummies are also attributed in order to verify if it has impact on IAI score. On models four and five the results are the same, the Materials and Financial sector are significant. The proposal of this study wasn't identifying why some sectors have more information asymmetry than others, but it was about identifying if IAI correctly measures its construct. Although, in agreement with literature, the financial sector is more regulated due to the probability of systematic risk if it collapses. Hence, it is expected that the disclosure policy of financial companies would be higher than non-financial, additionally the company with higher IAI score is ItaúUnibanco, one of the biggest Brazilian banks, and the first 7 companies with higher score are from the financial sector as well.

The ADR is shown by the literature as a metric that decreases information asymmetry. Lang, Lins, and Miller (2002) and Leuz (2003) discussed about cross-listing companies on Canada and United States the findings as interesting, it enhances analyst covering and forecast accuracy, also increases Tobin's Q. Those results are consistent with what literature shows as ways to decrease asymmetry. However, this paper finds no statistically significance on ADR dummy and cross-variables ADR with Bid and Ask Spread and ADR with Coverage.

## 5.5 General Findings

The proposal of this research was to build an accuracy information asymmetry model consistently with market analyst perception. The findings are consistent with literature in terms of variables signs and statistical significance, depending on the model chosen. It was noticed that sell side coverage, market-to-book ratio and stock volatility play an important role to determine IAI score, controlled for free float and traded value.

Also, it is interesting to find that the probability of informed trader combined with sell side coverage increases the perception of information asymmetry, showing companies exposure trough the sell side, it can have negative effects if they don't follow the best governance practices.

This work was concerned to run an efficacy model, testing and correcting for heteroskedasticity, collinearity, outliers, leverage and the correct measure of the variables, adapting Coverage, Traded Value and searching what literature provides for the best information asymmetry model. The results are satisfactory, the adjusted R squared of model five is 36,8%, which is a high number for cross-sectional models.

In this sense, this work proposes the model five teste as a model with secondary data reflecting the analysts' perception over companies' disclosure. That follows:

$$IAI_j = \alpha + \beta_1 COV_j + \beta_2 ERROR_j + \beta_3 VOL_j + \beta_4 GO_j + \beta_5 BaA_j + \beta_6 D_i + \beta_7 D_k + \epsilon$$

where:

*COV* = Proxy for Coverage, quantity of sell side analyst

*Error* = Proxy for the absolute difference between EPS median forecast and EPS realized

*Vol* = Volatility daily basis for stock

*GO* = Proxy for Growth Opportunities, it is the market-to-book ratio

*BaA* = Bid and Ask Spread

*D<sub>i</sub>* = Dummy for time

*D<sub>k</sub>* = Dummy for sector

Although there were some negative surprises in terms of data base (specifically on *PIN* and *Error* proxies) and statistical significance, research must continue, the IAI might have reach it maturity on  $t_4$ , more data enabled the usage of more sophisticated models as panel data, for example. In general, this works believes had been provided an innovation to literature, building an accuracy model based on market analyst perception of companies' disclosure, being a good metric for the information asymmetry research.

## 6. CONCLUDING REMARKS

The scope of this work was to create a model of information asymmetry with secondary data which reflects the analysts' perception over companies' disclosure. Their opinion was collected through a survey conducted among 2016-2018 and ranking companies into index named Information Asymmetry Index (IAI). To verify the adherence of the index, it was compared, through a pooled cross-section model, to five proxies selected from the literature review, which were divided in three groups: External Analysis, Internal Analysis and Market Microstructure. The findings showed that *Vol*, *Go* and *Cov* play an important role in the way to decrease information asymmetry, being statistically significant at almost all models ran indicating a high influence in IAI score.

Proxies from External Analysis was Coverage, Error and Volatility. The independent variable coverage (*Cov*) was used as the distance from sell side coverage of each company from the average coverage of all companies surveyed. Error is the difference between Bloomberg median EPS forecast from the EPS realized of the companies surveyed. Although, sell sides analysts are important to decrease market perception of information asymmetry, the independent variable wasn't significant at any model ran. The statistical insufficiency may be attributed to the high number of missing values. Volatility, as predicted on literature, decrease IAI score. The proxy was significant on all the models ran.

Internal Analyses was proxy of growth opportunities. Using market-to-book ratio the outcomes were significant in almost all the models ran, setting a positive association to the IAI score. Besides the critic surrounding this proxy, the findings apparently show that when investors buy stock, which are trading at a price higher than its book value is because they are aware of the risk assumed and are confident about the company's future.

On the Market microstructure, initially this work intends to use EHO model to infer the probability of informed trader on asset pricing, although the lack of data to all the period surveyed, PIN was replaced by Bid and Ask spread. Despite the theoretical assumption of higher spread, it is associated to higher levels of information asymmetry, this work findings were inconclusive about this proxy due to its lack of significance. PIN was used only at the last survey, conducted at November/18, but it also wasn't significant, it might have been if a longer period was available. However, a cross-variable between PIN and Coverage was used, the results showed a relevant finding

that sell side can increase information asymmetry if they understand that insider traders are acting on companies' asset pricing.

Dummies was used for time, sector and ADR issued. Time dummies show that, on average, recent facts are more important to determine the company's degree of information asymmetry. Although this assumption is in line with literature, the findings can be biased given the IAI methodology. At the beginning, all companies score 1500 points, over time disputes keep changing their score, carrying the past results to a survey form another, hence the outcome can be seen as IAI maturity instead of the recent fact importance. Sector dummies showed that Materials and Financial sectors make companies less asymmetric, although it wasn't the aim of this study to verify why some sectors have better disclosure practices than others, financial sectors are known by its regulation which necessarily enhances company's disclosure, hence it is satisfactory that the model captured this reality. ADR dummy didn't show significance to the model, instead of the theoretical assumption.

As shown above, information asymmetry affects companies over different forms, e.g. executive compensation, cost of capital, level of indebtedness, profitability, shareholder return, liquidity, control structure dividend policy, asset pricing and other, consequently different proxies can be used to measure this construct. Besides theme relevance, so far there isn't an agreement on how to correctly proxy it. The proposal of this work is to fill this gap providing an accurate model for information asymmetry.

The results were satisfactory, the model presents an adjusted R squared of 36,8% which is a great number for cross-sectional models and more suitable format to proxy the construct. It suggests the continuation of IAI survey for more data being available, enabling the usage of more sophisticated econometric model.

## REFERENCES

- ADAM, T.; GOYAL, V. K. The investment opportunity set and its proxy variables. **Journal of Financial Research**, v. 31, n. 1, p. 41–63, 2008.
- AKERLOF, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. **The Quarterly Journal of Economics**, 84(3), 488.
- ALBERS, P. C.; VRIES, H. d. Elo-rating as a tool in the sequential estimation of dominance strengths. **Animal Behaviour**, p. 489–495, 2001.
- ARMSTRONG, C. S.; CORE, J. E.; TAYLOR, D. J.; VERRECCHIA, R. E. When does information asymmetry affect the cost of capital. **Journal of Accounting Research**, v. 49, n. 1, p. 1–40, 2011.
- ASLAN, H. et al. The characteristics of informed trading: Implications for asset pricing. **Journal of Empirical Finance**, v. 18, n. 5, p. 782–801, 2011.
- BEYER, A.; COHEN, D. A.; LYS, T. Z.; WALTHER, B. R. The financial reporting environment: review of the recent literature. **Journal of accounting and economics**, v. 50, n. 2, p. 296–343, 2010.
- BOTOSAN, C. A. Disclosure level and the cost of equity capital. **Accounting review**, p. 323–349, 1997.
- BOSQUE, L. M. **Estimação da probabilidade de negociação privilegiada por meio de inferência bayesiana. 2016. 141 f.** 2016. Tese de Doutorado. Dissertação (Mestrado em Administração: Finanças). Universidade de Brasília. Brasília.
- BRENNAN, M. J.; TAMAROWSKI, C. Investor relations, liquidity, and stock prices. **Journal of Applied Corporate Finance**, v. 12, n. 4, p. 26–37, 2000.
- BROWN, L. D.; RICHARDSON, G. D.; SCHWAGER, S. J. An information interpretation of financial analyst superiority in forecasting earnings. **Journal of Accounting Research**, p. 49–67, 1987.
- BUSHEE, B. J.; MILLER, G. S. Investor relations, firm visibility, and investor following. **The Accounting Review**, v. 87, n. 3, p. 867–897, 2012.
- ÇELİK, Duygu; TINIÇ, Murat. InfoTrad: An R package for estimating the probability of informed trading. **Available at SSRN 2889323**, 2017.
- CLARKE, J., & SHASTRI, K. (2001). On Information Asymmetry Metrics. **Ssrn, 1708**. <https://doi.org/10.2139/ssrn.251938>
- DIAMOND, D. W.; VERRECCHIA, R. E. Disclosure, liquidity, and the cost of capital. **The journal of Finance**, v. 46, n. 4, p. 1325–1359, 1991.
- DIERKENS, Nathalie. Information asymmetry and equity issues. **Journal of Financial and Quantitative Analysis**, v. 26, n. 2, p. 181–199, 1991.

- DUARTE, J. et al. Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital. **Journal of Financial Economics**, v. 87, n. 1, p. 24-44, 2008.
- EASLEY, D., ENGLE, R. F., O'HARA, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed trades. **Journal of Financial Econometrics**, 6(2), 171–207.
- EASLEY, D.; HVIDKJAER, S.; O'HARA, M. Is information risk a determinant of asset returns. **The journal of finance**, v. 57, n. 5, p. 2185–2221, 2002.
- EASLEY, D.; HVIDKJAER, S.; O'HARA, M. Factoring information into returns. **Journal of Financial and Quantitative Analysis**, v. 45, n. 2, p. 293-309, 2010.
- EASLEY, D.; KIEFER, N. M.; O'HARA, M. The information content of the trading process. **Journal of Empirical Finance**, v. 4, n. 2-3, p. 159–186, 1997.
- EASLEY, D.; O'HARA, M. Price, trade size, and information in securities markets. **Journal of Financial economics**, v. 19, n. 1, p. 69–90, 1987.
- EASLEY, D.; O'HARA, M. Information and the cost of capital. **The Journal of Finance**, v. 59, n. 4, p. 1553–1583, 2004.
- ELO, A. New USCF rating system. *Chess Life*, v. 16, p. 160–161, 1961
- Fama, E. F. (1969). American Finance Association Efficient Capital Markets : A Review of Theory and Empirical Work. **The Journal of Finance**, 25(2), 28–30.
- GLOSTEN, L. R.; HARRIS, L. E. Estimating the components of the bid/ask spread. **Journal of financial Economics**, v. 21, n. 1, p. 123–142, 1988.
- HARRIS, M.; RAVIV, A. The theory of capital structure. **The Journal of Finance**, v. 46, n. 1, p. 297–355, 1991.
- HE, W. P.; LEPONE, A.; LEUNG, H. Information asymmetry and the cost of equity capital. **International Review of Economics & Finance**, v. 27, p. 611–620, 2013.
- HEALY, P. M.; PALEPU, K. G. Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. **Journal of Accounting and Economics**, v. 31, n. 1-3, p. 405–440, 2001.
- HUDDART, S. J.; KE, Bin. Information asymmetry and cross-sectional variation in insider trading. **Contemporary Accounting Research**, v. 24, n. 1, p. 195-232, 2007.
- HVIDKJAER, S. Probability of Informed Trading. **Encyclopedia of Quantitative Finance**, 2010.
- KYLE, A. S. Continuous auctions and insider trading. **Econometrica: Journal of the Econometric Society**, p. 1315–1335, 1985.



- LAMBERT, R. A.; LEUZ, C.; VERRECCHIA, R. E. Information asymmetry, information precision, and the cost of capital. **Review of Finance**, v. 16, n. 1, p. 1–29, 2011.
- LANG, Mark H.; LINS, Karl V.; MILLER, Darius P. ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value?. **Journal of Accounting Research**, v. 41, n. 2, p. 317-345, 2003.
- LANG, M.; LUNDHOLM, R. Cross-sectional determinants of analyst ratings of corporate disclosures. **Journal of Accounting Research**, v. 31, n. 2, p. 246–271, 1993.
- LI, Kevin K.; YOU, Haifeng. What is the value of sell-side analysts? Evidence from coverage initiations and terminations. **Journal of Accounting and Economics**, v. 60, n. 2-3, p. 141-160, 2015.
- H.W. W. Lin and W. C. Ke. A computing bias in estimating the probability of informed trading. **Journal of Financial Markets**, 14(4):625-640, 2011.
- LEUZ, Christian. Discussion of ADRs, Analysts, and Accuracy: Does Cross-Listing in the United States Improve a Firm's Information Environment and Increase Market Value?. **Journal of Accounting Research**, v. 41, n. 2, p. 347-362, 2003.
- MARTINS, Orleans Silva; PAULO, Edilson. Assimetria de informação na negociação de ações, características econômico-financeiras e governança corporativa no mercado acionário brasileiro. **Revista Contabilidade & Finanças**, v. 25, n. 64, p. 33-45, 2014.
- MCLAUGHLIN, Robyn; SAFIEDDINE, Assem; VASUDEVAN, Gopala K. The information content of corporate offerings of seasoned securities: An empirical analysis. **Financial Management**, p. 31-45, 1998.
- MODIGLIANI, F.; MILLER, M. H. The cost of capital, corporation finance and the theory of investment. **The American**, v. 1, p. 3, 1958.
- MYERS, S. C.; MAJLUF, N. S. Corporate financing and investment decisions when firms have information that investors do not have. **Journal of Financial Economics**, v. 13, n. 2, p. 187–221, 1984.
- PENMAN, Stephen H. The articulation of price-earnings ratios and market-to-book ratios and the evaluation of growth. **Journal of accounting research**, v. 34, n. 2, p. 235-259, 1996.
- PERLIN, Marcelo; RAMOS, Henrique. GetHFData: AR package for downloading and aggregating high frequency trading data from Bovespa. **Available at SSRN 2824058**, 2016.
- RABELO, F. M.; VASCONCELOS, F. C. Corporate governance in Brazil. **Journal of Business Ethics**, v. 37, n. 3, p. 321–335, 2002.

- RANI, A. J.; KARAMANOU, I. Regulation fair disclosure, analyst following, and analyst forecast dispersion. **Accounting Horizons**, v. 17, n. 1, p. 15–29, 2003.
- ROSS, S. A. The determination of financial structure: the incentive-signalling approach. **The Bell Journal of Economics**, p. 23–40, 1977.
- SCHIEMANN, F., RICHTER, K., & GÜNTHER, T. (2010). Voluntary disclosure of intellectual capital items in roadshows and investor conferences: An empirical analysis of DAX30-companies. **Zeitschrift Fur Planung Und Unternehmenssteuerung**, 21(3), 255–275.
- SHIN, H.-H.; STULZ, R. M. Firm value, risk, and growth opportunities. **National Bureau Of Economic Research**, 2000.
- SMITH JR, Clifford W.; WATTS, Ross L. The investment opportunity set and corporate financing, dividend, and compensation policies. **Journal of financial Economics**, v. 32, n. 3, p. 263-292, 1992.
- SONAS, Jeff. The sonas rating formula-better than elo. **Chess Base Magazine Online**, 2002.
- STIGLITZ, J. E. Knowledge as a global public good. **Global Public Goods: International Cooperation in The 21st Century**, v. 308, p. 308–25, 1999.
- STIGLITZ, J. E. The Revolution of Information Economics: the past and the future. **National Bureau of Economic Research**, 2017.
- STOLL, H. R.; CURLEY, A. J. Small business and the new issues market for equities. **Journal of Financial and Quantitative Analysis**, v. 5, n. 3, p. 309–322, 1970.
- THOMAS, S. Firm diversification and asymmetric information: evidence from analysts' forecasts and earnings announcements. **Journal of Financial Economics**, v. 64, n. 3, p. 373–396, 2002
- USATEGUI, J. M. Economía de la Información. Servicio **Editorial de la Universidad del País Vasco**, 2000.
- VERRECCHIA, R. E. Discretionary disclosure. **Journal of Accounting and Economics**, v. 5, p. 179–194, 1983.
- WANG, J. A model of intertemporal asset prices under asymmetric information. **The Review of Economic Studies**, v. 60, n. 2, p. 249–282, 1993.

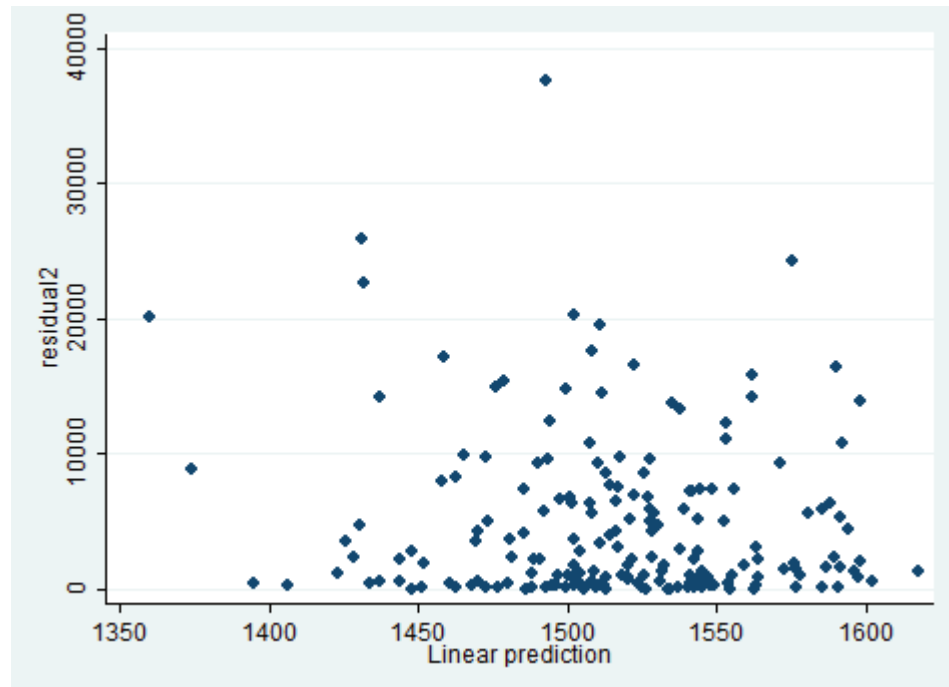
## APPENDIX A – ECONOMETRIC TESTS

### A.1 Multicollinearity Test – Variance Inflation Test

Variable	VIF	1/VIF
BaA	1.49	0.672267
Cov	1.38	0.726838
Intradedvalue	1.23	0.811966
Vol	1.13	0.883927
Error	1.09	0.913901
ff	1.07	0.931700
Go	1.06	0.941640
Mean	VIF	1.21

### A.2 Homoscedasticity Tests

Graph test for Heteroskedasticity



Source: The Author

### Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of IAI

chi2(1) = 1.98

Prob > chi2 = 0.1595

**White test**

Ho: Constant variance

White's general test statistic : 54.38058 Chi-sq(35) P-value = .0194



## APPENDIX B – DISCLOSURE WEBSITE

← → 🔍 Não seguro | disclosureindex.com/br/dupla.asp?insc=261 ☆ 📄 📱 📶 🔋 🌐 ⋮

Disputa 1 de 10  
Qual destas empresas apresenta melhor nível de disclosure?

**Considere disclosure como:** "Comunicação de todas as informações positivas e negativas sobre a empresa. O seu objetivo é permitir aos credores e investidores formarem opinião sobre a situação financeira da empresa."

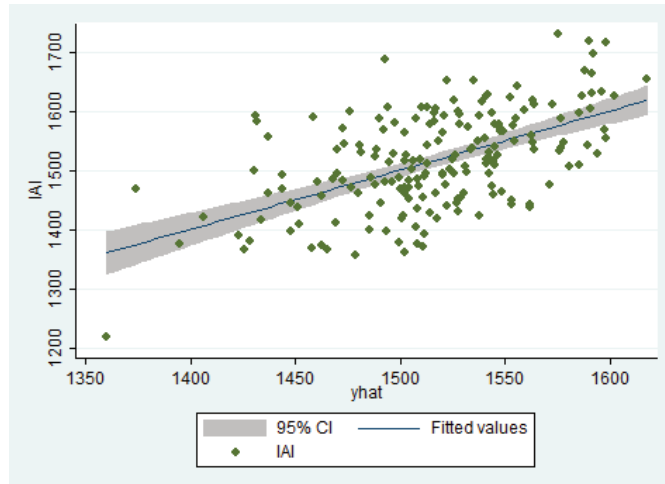
**Importante:** O que nos importa é sua percepção inicial. Alguns segundos são suficientes para sua escolha.

## APPENDIX C – PLOTS

### C.1 – IAI PLOT

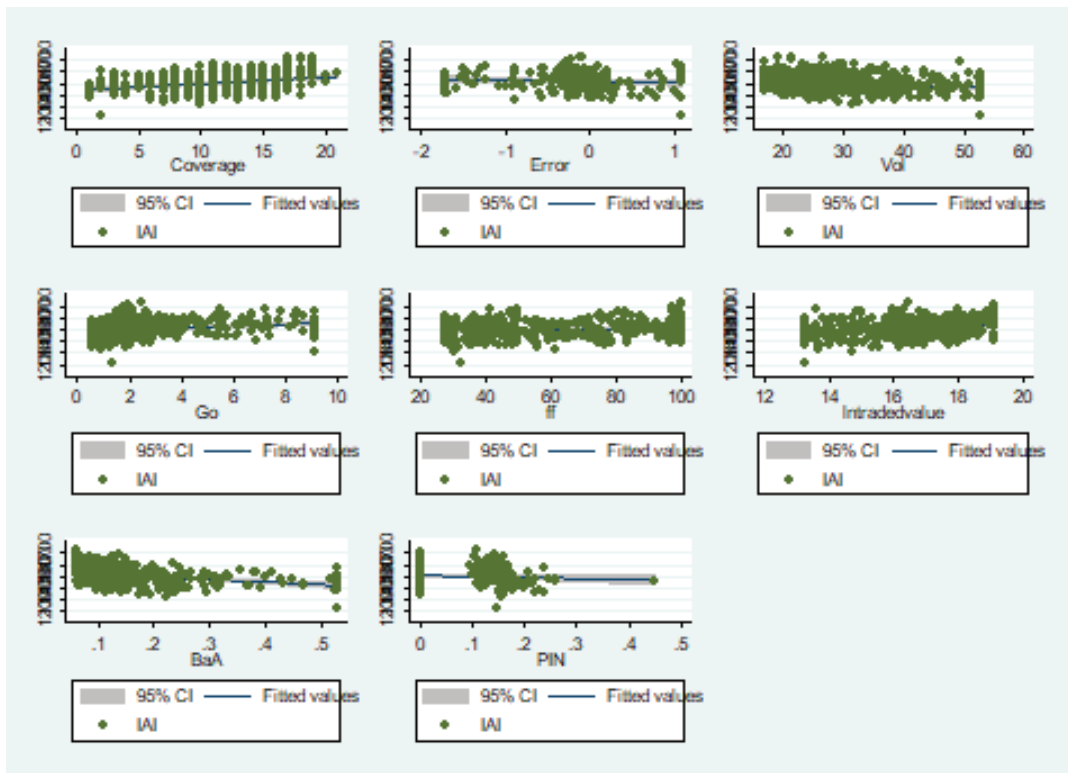
Graph IAI, yhat



Source: The Author

### C.1 – INDEPENDENT VARIABLE PLOT

Combine Graph IAI, independent Variable



Source: The Author