FEDERAL UNIVERSITY OF MINAS GERAIS
FACULTY OF ECONOMIC SCIENCES
DEPARTMENT OF ADMINISTRATIVE SCIENCES
CENTRE OF POSTGRADUATE AND RESEARCH IN ADMINISTRATION

CAROLINA MAGDA DA SILVA ROMA

IDIOSYNCRATIC VOLATILITY: AN ANALYSIS OF AGGREGATE AND INDIVIDUAL EFFECTS

BELO HORIZONTE

2017

CAROLINA MAGDA DA SILVA ROMA

IDIOSYNCRATIC VOLATILITY: AN ANALYSIS OF AGGREGATE AND INDIVIDUAL EFFECTS

Thesis submitted to the Centre of Postgraduate and Research in Administration of the Federal University of Minas Gerais as a partial requirement of the Doctorate in Administration Program.

Supervisor: Dr. Robert Aldo Iquiapaza

Co-supervisor: Dr. Hudson Fernandes Amaral

Belo Horizonte

2017

Ficha Catalográfica

Roma, Carolina Magda da Silva.
R756i Idiosyncratic volatility [manuscrito] : an analysis of aggregate
2017 and individual effects / Carolina Magda da Silva Roma . – 2017.
144 f.: il., gráfs. e tabs.

Supervisor : Robert Aldo Iquiapaza
Co-supervisor : Hudson Fernandes Amaral.
Tese (doutorado) - Universidade Federal de Minas Gerais, Centro de Pós-Graduação e Pesquisas em Administração.
Inclui bibliografia (f. 135-144)

1. Mercado – Teses. 2. Administração – Teses. I. Iquiapaza, Robert Aldo. II. Amaral, Hudson Fernandes. III. Universidade Federal de Minas Gerais. Centro de Pós-Graduação e Pesquisas em Administração. IV. Título

CDD: 658

Elaborada pela Biblioteca da FACE/UFMG. – NMM/086/2017

Universidade Federal de Minas Gerais
Faculdade de Ciências Econômicas
Departamento de Ciências Administrativas
Centro de Pós-Graduação e Pesquisas em Administração

ATA DA DEFESA DE TESE DE DOUTORADO EM ADMINISTRAÇÃO da Senhora **CAROLINA MAGDA DA SILVA ROMA**, REGISTRO Nº 177/2017. No dia 03 de maio de 2017, às 14:00 horas, reuniu-se na Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais - UFMG, a Comissão Examinadora de Tese, indicada pelo Colegiado do Centro de Pós-Graduação e Pesquisas em Administração do CEPEAD, em 24 de abril de 2017, para julgar o trabalho final intitulado **"Idiosyncratic Volatility: an analysis of aggregate and individual effects"**, requisito para a obtenção do **Grau de Doutor em Administração**, linha de pesquisa: **Finanças.** Abrindo a sessão, o Senhor Presidente da Comissão, Prof. Dr. Robert Aldo Iquiapaza Coaguila, após dar conhecimento aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa da candidata. Logo após, a Comissão se reuniu sem a presença da candidata e do público, para julgamento e expedição do seguinte resultado final:

(X) APROVAÇÃO;

( ) APROVAÇÃO CONDICIONADA A SATISFAÇÃO DAS EXIGÊNCIAS CONSTANTES NO VERSO DESTA FOLHA, NO PRAZO FIXADO PELA BANCA EXAMINADORA (NÃO SUPERIOR A 90 NOVENTA DIAS);

( ) REPROVAÇÃO.

O resultado final foi comunicado publicamente à candidata pelo Senhor Presidente da Comissão. Nada mais havendo a tratar, o Senhor Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 03 de maio de 2017.

NOMES	ASSINATURAS
Prof. Dr. Robert Aldo Iquiapaza Coaguila ORIENTADOR (CEPEAD/UFMG)	....
Prof. Dr. Hudson Fernandes Amaral (CEPEAD/UFMG)	....
Prof. Dr. Aureliano Angel Bressan (CEPEAD/UFMG)	....
Prof. Dr. Wagner Moura Lamounier (CEPCON/UFMG)	....
Prof. Dr. Clayton Peixoto Goulart (Banco Central do Brasil)	....
Prof. Dr. Herbert Kimura (Universidade de Brasília/ DF)	....

To God for the gift of life

and to my lovely family

for everything they represent to me

Ackowledgements

First of all, I want to thank God for the opportunity of life and the chance of standing here, right now, concluding this cycle that started four years ago.

I want to thank my family for all their support in this long journey. I want to thank my father, Paulo, who God has taken from us physically some years ago, but has continued existing in our actions, memories, and hearts. He was a great man who did his best for our family and never disappointed us. He was there to take us to school and waited for us to come home with a beautiful smile on his face and so much love in his heart. He did everything he could for us. His existence changed mine and made me who I am. I will always love you, daddy!

To my mother, Magda, for being my wonderful and beautiful mother, for continuously supporting me and motivating me to follow my dreams and my heart, for always believing in me and being close to help me in anything I needed, for the chance of being more than I ever could have imagined to be (and you were there with me!). You have always given me the wings to fly and I am here because of you, because you dreamed that I could accomplish this goal, because you love me. Since I was a child, your love guided me and made it possible for me to stand here today and write down this thesis. I am very proud of being your youngest daughter. I love you, mother.

To my great sister Paula, who has unconditionally been with me, guided me through her own experiences, supported me day after day, for the patience to teach me and help me whenever I needed, day or night, because she has never left me alone, she has never let me down. I decided to study and do my PhD because of her inspiration and desire to learn and go deep further in whatever she decides to do. Every single weekend that I needed to come to UFMG, she was there with me. I have no words to say how much you influenced who I am, where I am now. Sister, you are my personal super hero, my older sister, my best friend, and the one who, well, translates what love is. I really do not know where I would be without you. I love you sister.

To my fiancée for loving me. He has always been such a wonderful partner to share my life with and has taught me how to be a better person, and what life can be. I really admire him and our family and I are really proud of the man he has become. God knows how much I am thankful for your life and for the chance of sharing beautiful moments with you and our family. You brought a new history to me and you have always made me smile. Undoubtedly, you made this journey much easier with your support and motivation to be happy, to work harder, and to never give up. You have been an inspiration to me and guided my choices. I do not know if I would be here without you. Thank you, darling. I love you.

I also want to express my gratitude to my advisor, Professor Robert Aldo Iquiapaza. He has provided me the chances to learn Finance in a much deeper way during this course with very insightful ideas. His door was always open to me and he was always there sharing his knowledge with me. He was always ready to discuss any topic I had ever wanted to study. He gave me all the opportunities he could and trusted me. I am really proud to have shared four years of my life with him. Thank you, professor.

To my co-advisor, Professor Hudson Fernandes Amaral, for giving me the chance of being one of his students and for always receiving me in a warm way in his room and providing me with rich moments of discussion and knowledge. What I learned from you made it possible for me to write down this thesis today. Thank you, professor.

To my co-advisor, Professor Roméo Tedóngap, who advised me since my ‘sandwich doctorate’ in Stockholm. He gave me the chance of being part of his institution and that was one of the greatest experiences in my life. His school is a place I could never have imagined before. Through his acceptance to supervise me, I had the chance to attend classes, seminars, and courses and to redirect my work to a new literature. I have no words to describe how thankful I am for this big chance and for trusting me. Thank you, professor.

My deep acknowledgments are also given to all the professors I had contact during these four years in CEPEAD. Everything you taught me contributed to my research and opened my mind to a new world. I am really thankful for the opportunity of having you as my professors. I also could not forget to mention that I am thankful to Professor Felipe Paiva, from CEFET/MG, who gave me access to part of my data and was always ready to contribute to my research. Thank you, professors.

To my friends in CEPEAD for making my life very happy and for making me smile every single day I was there. I cannot name all of you, so I am not going to try. But, you were very good friends who shared your lives with me, and a totally new learning opportunity. Thank you to all of you, my dear friends.

Thank you to Érika, Vera, Luciana, and Ana Paula who work in the CEPEAD for all assistance during this period.

Finally, I want to say thank you to FAPEMIG and CAPES for their financial support with my scholarship here in Brazil and during my sandwich period, respectively, because this research was also only possible with their support.

"I would rather have questions that can't be answered than answers that can't be questioned."

Richard Feynman

ABSTRACT

This work is focused on the study of idiosyncratic risk at the aggregate and firm-level and its effect on returns. For that analysis we use a model-free measure, namely the cross-sectional variance (CSV), easily obtained at any frequency, as proposed in Garcia, Mantilla-García, and Martellini (2014). This measure was employed in the U.S. and Brazilian stock markets using equal- and value-weighted market returns, at daily and monthly frequencies. For U.S. stocks, first the main findings in the aforementioned authors' work were replicated and their sample was extended until December 2014. The results relative to these estimations imply that aggregate idiosyncratic volatility is still positive and significant in predicting equal- and value-weighted market returns. In a second moment, credit ratings of firms rated by Standard & Poor's were used to construct a CSV measure according to the firm's credit quality into investment grade, non-investment grade and all rated firms. The analysis relying on CSV by levels of ratings pointed out to the relevance of monthly value-weighted measures in predicting market returns in the whole sample. Using Brazilian stocks from January 2000 to June 2016, the behavior of CSV was investigated as a proxy to idiosyncratic variance over time and its predictive power to forecast market returns throughout the whole timespan, subperiods, up and down markets, expansion and recession periods as well testing its power with the inclusion of market variance, investor sentiment, aggregate dividend yield and expected and unexpected measures of illiquidity. It was also analyzed whether that measure can predict risk factors omitted from the CAPM model as presented in Carhart (1997) and Fama and French (2015) five-factor model. Overall, these findings suggest that the firm-specific risk aggregated over the stocks in the sample is not able to robustly predict market returns or risk factors. Using credit ratings of Brazilian firms showed to be very limited due to the number of rated companies, therefore a new CSV based on ratings or the stock's market capitalization (size) were constructed; however, these approaches also do not corroborate the idea that aggregate idiosyncratic variance is useful in predicting market returns. After that, attention was focused on understanding the cross-sectional effects of expected idiosyncratic volatility and expected return using an EGARCH model developed by Nelson (1991) and a Skewed – Generalized Error Distribution (Skew-GED). Portfolios were formed based only on expected idiosyncratic risk and controlling for other characteristics which have been shown to affect the stock return, specifically size, book-to-market, momentum and reversal return applying Ang *et al.*'s (2006) methodology to account for these variables in a double sorting procedure. The abnormal returns generated by these portfolios, in general, are not statistically significant when controlling for other characteristics. Fama and MacBeth (1973) cross-sectional regressions to evaluate the relation between expected returns and expected idiosyncratic volatility controlling for portfolio beta, market capitalization, book-to-market, turnover, momentum, coefficient of variation of turnover, and lagged return were run. Using the three different models consistently suggest that expected idiosyncratic volatility is not related to expected returns when the forward-looking return observation is not included in the estimations supporting Fink, Fink and He (2012) results.

Keywords: CSV; EGARCH; Expected Return

RESUMO

Este trabalho tem como foco o estudo do risco idiossincrático nos níveis agregado e da firma e seu efeito sobre os retornos. Para essa análise, usamos uma medida independente de modelos, nomeadamente a variância cross-sectional (cross-sectional variance, CSV), facilmente obtida em qualquer frequência, como proposto em Garcia, Mantilla-García e Martellini (2014). Essa medida foi empregada nos mercados de ações dos Estados Unidos e Brasil usando retornos de mercado igualmente ponderados e por capitalização de mercado, nas frequências diária e mensal. Para ações dos Estados Unidos, primeiro os resultados principais do trabalho dos autores previamente mencionados foram replicados e sua amostra foi expandida até Dezembro 2014. Os resultados relativos a essas estimações implicam que a volatilidade idiossincrática agregada é ainda positiva e significante para prever retornos de mercado igualmente ponderados e por capitalização de mercado. Em um segundo momento, ratings de crédito de firmas rateadas pela Standard & Poor's foram usados para construir a medida CSV de acordo com a qualidade de crédito da firma em grau de investimento, especulativo e todas as firmas rateadas. As análises sobre a CSV por níveis de ratings apontaram a relevância das medidas mensais com capitalização de mercado para prever retornos de mercado na amostra completa. Usando ações brasileiras de janeiro de 2000 até junho de 2016, o comportamento da CSV foi investigado como uma proxy para a variância idiossincrática ao longo do tempo e seu poder preditivo para prever retornos de mercado durante todo o período de tempo, subperíodos, mercados em alta e baixa, períodos de expansão e recessão, como também seu poder foi testado com a inclusão da variância de mercado, sentimento do investidor, dividend yield agregado e medidas de iliquidez esperada e inesperada. Também foi analisado se a medida pode prever fatores de risco omitidos do modelo CAPM, como apresentado em Carhart (1997) e no modelo de cinco fatores de Fama e French (2015). De maneira geral, esses resultados sugerem que o risco específico da firma agregado sobre as ações na amostra não é capaz de robustamente prever retornos de mercado ou fatores de risco. Usando ratings de crédito de firmas brasileiras mostrou-se muito limitado devido ao número de companhias avaliadas, assim uma nova CSV baseada em ratings ou na capitalização de mercado da ação (tamanho) foram construídas; contudo, essas abordagens também não corroboram a ideia de que a variância idiossincrática agregada é útil em prever retornos de mercado. Depois disso, a atenção foi focada no entendimento dos efeitos cross-sectional da volatilidade idiossincrática esperada e retornos esperados usando o modelo EGARCH desenvolvido por Nelson (1991) e a Distribuição Generalizada do Erro – Assimétrica (Skewed – Generalized Error Distribution, Skew-GED). Portfólios foram formados baseados somente no risco idiossincrático esperado e controlando por outras características que têm mostrado afetar o retorno da ação, especificamente tamanho, valor patrimonial/valor de mercado, momento e reversão do retorno aplicando a metodologia de Ang *et al.* (2006) para considerar essas variáveis em um procedimento de sorteio bivariado. Os retornos anormais gerados por esses portfólios, em geral, não são estatisticamente significantes quando controlando por outras características. Regressões cross-section de Fama e MacBeth (1973) para avaliar a relação entre retornos esperados e volatilidade idiossincrática esperada controlando pelo beta da carteira, capitalização de mercado, valor patrimonial/valor de mercado, turnover, momento, coeficiente de variação do turnover e retorno defasado foram realizadas. Foram usados três diferentes modelos que consistentemente sugerem que a volatilidade idiossincrática esperada não é relacionada aos retornos esperados quando o retorno futuro não é incluído nas estimações suportando os resultados de Fink, Fink e He (2012).

Palavras-chave: CSV; EGARCH; Retorno Esperado

List of Figures

List of Tables

Table of Contents

1 INTRODUCTION

1.1 Background

The classical financial theory in the well-known Capital Asset Pricing Model (CAPM) developed by Sharpe (1964), Lintner (1965a) and Mossin (1966) states that in equilibrium the only parameter that matters to the investor is the systematic risk given by the exposure the asset's return has to a single risk factor, the beta. However, idiosyncratic risk is the part of total risk that is not explained by systematic factors and, hence, reflects the firm-level specific risk.

Supposing that enough diversification is made by investors and a correct model specification is given by CAPM or the three-factor model by Fama and French (1993, hereafter FF-3), it is not expected that any important role would be found for the residuals, i.e. the part not explained of the model (BALI, CAKICI, LEVY, 2008). As mentioned by Bali, Cakici and Whitelaw (2011), diversification has a crucial importance in this model as it reduces the total risk, specifically the firm-specific risk, which has no role in this asset pricing theory. However, Mitton and Vorkink (2007) and Goetzmann and Kumar (2008) point out that investors do not fully diversify their portfolios.

These results are also corroborated by the work of Campbell *et al.* (2001), who decompose stock volatility at the firm, industry and market level during the period 1962-1997. The authors document that the volatility at the industry and market levels have been stable over time in opposition to the upward trend verified in the average idiosyncratic firm volatility. In other words, it means that the correlation among stocks has decreased and it translates into an increase in the number of stocks held by investors to achieve diversification.

Xu and Malkiel (2003) evidence a rise in the idiosyncratic volatility and link it to the increase in institutional ownership and high growth. In contrast, Brandt *et al.* (2010) reveal that the increase in idiosyncratic volatility identified by Campbell *et al.* (2001) cannot be seen as a time trend but rather an episodic phenomenon as the evidences presented in their papers show that by 2003 idiosyncratic volatility had already decreased to its level before 1990.

Several propositions regarding the computation of idiosyncratic variance (volatility) have been made in the literature in recent years. Therefore, if investors do not fully diversify their portfolios, idiosyncratic risk may be a risk factor that affects the investor portfolio and if this is true then investors would require a risk premia for dealing with it. Additionally, it is commonly understood that higher risk should generate higher return, then if idiosyncratic volatility is a priced factor, it is expected to have a positive relation with future returns. Indeed, theories as in Merton (1987) and Malkiel and Xu (2002) indicate a positive relation between the unsystematic risk and expected returns.

In contrast, Ang *et al.* (2006) analyzed the U.S stock market, and later 23 countries in Ang *et al.* (2009), revealing a negative relation between realized idiosyncratic volatility (*IVOL*) at the firm-level and expected stock returns. Specifically, Ang *et al.* (2006) document that a strategy that longs stocks with higher idiosyncratic volatility in the previous month and shorts those with the lowest idiosyncratic volatility generates a significant monthly average raw return of -1.06%, implying that stocks with higher realized idiosyncratic volatility earn lower returns, a phenomenon known in the literature as the “idiosyncratic risk puzzle”, which corresponds to an anomalous result attracting the attention of many researches.

This controversial result has been intensively investigated by other academics making use of American data, as in Huang *et al.* (2010), as well in other countries around the world, as in Brockman, Schutte and Yu (2009). In turn, distinct reasons have emerged when trying to explain this conflicting result. In this sense, Bali and Cakici’s (2008) analysis show that Ang *et al.*’s (2006) results are driven by small and illiquid stocks and sensitive to the method employed.

Other academics have uncovered a positive and significant relation between idiosyncratic volatility and expected returns. Fu (2009) advocates that Ang *et al.*’s (2006, 2009) measure of idiosyncratic volatility, computed as the standard deviation of residuals from a within-month regression of daily stock returns against risk factors in the Fama and French (1993) model, is not a correct proxy for the contemporaneous idiosyncratic risk as the author has shown that the idiosyncratic volatility given in that way does not follow a random walk process.

Instead of using lagged idiosyncratic volatility, Fu (2009) suggests calculating it from a conditional volatility perspective using the Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH) model of Nelson (1991), though using a forward-looking observation in the forecasted idiosyncratic risk estimation. This measure is known as the expected idiosyncratic volatility. Moreover, the author finds that the return reversal helps to explain the negative relation between future returns and lagged IVOL idiosyncratic volatility as the latter becomes non-significant when it is included in the regressions. Huang *et al.* (2010) find similar results in favor of return reversal.

Guo, Kassa and Ferguson (2014) claim that Fu's (2009) results suffer from a look-ahead bias given the incorporation of one data point (one forward-looking observation), i.e. the author uses data up to time *t* to forecast the idiosyncratic volatility in time *t.* In response to this new evidecence, Fu (2010) argues that their implementation has two major drawbacks: first, the inclusion of forecasted idiosyncratic volatilities, which does not necessarily have a status reporting that the model in fact achieved convergence; second, in their settings the maximum number of iterations is 100, which Fu (2010) points out to be very restrictive, suggesting that 500 (his benchmark setting) leads to improvements of the estimated models.

However, Fink, Fink and He (2012) show that if no forward-looking return observation is included to estimate the conditional idiosyncratic risk, then it has no significant relation with expected returns following both Fu's (2009) or Ang *et al.*'s (2006) procedures, but when forward-looking observation is added then the positive relation is recovered, as in Fu (2009). Therefore, it seems important to take a better look at the relation between expected idiosyncratic risk and expected returns at the firm-level taking into account a truly out-of-sample experiment to verify if Fink, Fink and He (2012) can be supported through an emerging country perspective.

At the aggregate level, there is also a lively debate. Goyal and Santa-Clara (2003) affirm that equal-weighted total variance (mostly driven by idiosyncratic variance) matters to predict value-weighted market returns during the period 1962 to 1999. Bali *et al.* (2005) argue that Goyal and Santa-Clara's (2003) results do not hold in an extended sample and document that there is no significant relation between value-weighted stock volatility and value-weighted market returns.

Wei and Zhang (2005) revisit Goyal and Santa-Clara's (2003) findings through additional data covering the period up to december 2002 and do not find it to be robust accross different subsamples and when considering the extended period. As opposed to the focus on the relation between value-weighted average return and equally-weighted average volatility, Wei and Zhang (2005) report that the equally-weighted average return and the equally-weighted average volatility or cross-sectional variance have a positive and significant relation, though it does not generate economic gains either.

Guo and Savickas (2006) highlight that the reason for some previous studies having failed in finding a posivite tradeoff between risk and return is due to an omitted variable problem. When market volatility and idiossyncratic risk are taken in the analysis jointly, the authors obtain that both are significant, though with opposite signals, in which the former is positive and the latter is negative. Therefore, some authors report a positive, non-significant and negative relation when studying the forecasting power of aggregate idiosyncratic variance to market returns.

Notwithstanding, another key point is the analysis of whether the credit quality of the firms is somehow related to the companies' specific risk. Therefore, the link between credit ratings and idiosyncratic risk is still a question that needs more investigation. Credit ratings may matter as they represent one way to classify firms according to their levels of risk. Fitch, Moody's and Standard & Poor's are the three main agencies which classify companies according to the perspective of meeting its financial obligations as they come due. The Long-Term Issuer Credit Ratings provided by Standard & Poor's represent a "forward-looking opinion about an obligor's overall creditworthiness" (STANDARD & POOR'S**,** 2016) and is not related to any specific financial obligation or financial program. Some of the previous works using credit ratings offer an understanding of how they are linked with the evaluation of the stock's performance, information asymmetry, corporate governance, dispersion in analyst's earnings forecast and in the properties of the credit ratings.

In this sense, Avramov *et al.* (2007) show the relation that exists between credit ratings and the momentum profit strategy designed by Jegadeesh and Titman (1993), through which low-grade firms can have momentum profit while the same is not true for high-grade firms. Avramov *et al.* (2009) also study why high-rated firms earn higher returns than low-rated firms. They find that the negative relation in credit risk and retuns preponderate only in the period closer to the credit rating downgrade (3 months before and after) and is related to the worst-rated firms in financial distress. In addition, in stable or improving rating periods, they find no difference in the returns between high credit quality firms and worse credit quality firms. Therefore, it may be useful to understand the influence of idiosyncratic risk on expected returns by levels of credit ratings as well as stable and deteroritating ratings periods.

Ashbaugh-Skaife, Collins and LaFond (2006) investigate whether firms with better corporate governance can obtain higher credit ratings in comparison with those firms with weaker governance practice and find out that the firm's governance level influences its overall credit rating. Tang (2009) exploits the refinement made by Moody's in the credit rating classification with the introduction of levels of specification as an additional distinction between the credit quality of the firms, resulting in nineteen categories. The author argues that the refinement affected investors in the sense that they could make better inferences about the firm, while the firm is also influenced in the capital market access through the reduction in the cost of borrowing for those with better refined ratings; it also impacts the firm's investment decisions. Hence, investors can use the information contained in the assigned credit rating of a firm to make better decisions.

Given that the literature has made attempts to study different dimensions related to credit ratings, it is interesting to note that the investigation about its relation to aggregate idiosyncratic variance estimators requires a deeper analysis. As pointed by Avramov *et al.* (2007), the credit ratings follow the business cycle. Then, it is also of importance to understand the behavior of the idiosyncratic risk of stocks segmented according to the firm's credit quality.

In this sense, this thesis aims to analyse if aggregate or individual idiosyncratic risks are related to returns. To achieve this goal, its methods are threefold. Firstly, it presents a comprehensive understanding of the aggregate idiosyncratic variance for which the benefit from a new measure of cross-sectional variance developed by Garcia, Mantilla-García, and Martellini (2014) is taken, which has two main attributes: it is a model-free measure and can be computed at any frequency, thereby overcoming some limitations of previous work that used mainly monthly data; and it is highly correlated with idiosyncratic risk using U.S. and Brazilian stock market.

The role of market variance and investor sentiment is also considered when studying whether an association between aggregate idiosyncratic risk and market returns exists. Traditional theory on asset pricing claims that only systematic risk matters and competition among rational investors, who act in order to optimize their portfolios, will induce equilibrium where price represents the discounted value of expected cash-flows (BAKER, WURGLER, 2006).

However, recent theories and empirical work have been done trying to understand the impact of investor attitudes or sentiment on the cross-sectional and time series behavior of return and risk. Baker and Wurgler (2006) point out that investor sentiment can be defined as a propensity to speculate or optimism and pessimism about stocks. Brown and Cliff (2005) use consumer confidence to proxy for investor sentiment and point out a negative relation between market returns and investor sentiment for the U.S. stock market. Schmeling (2009) finds that market returns are negatively related to investor sentiment measured by the consumer confidence using an international sample of developed countries. Then, it becomes important to understand whether and how investor sentiment affects the relation between aggregate idiosyncratic risk and market returns.

The analysis of cross-sectional skewness and kurtosis are useful to predict market returns and whether aggregate idiosyncratic variance helps predict risk factors omitted from the CAPM model are investigated as well. In this sense, Garcia, Mantilla-García, and Martellini (2014) study the effect of asymmetry in aggregate idiosyncratic variance by splitting the returns according to their position in relation to the market returns, i.e. the returns to the right and to the left of the market returns. Their results reflect an asymmetric behavior in the data pattern and use a measure of cross-sectional skewness to predict market returns. Herein, the analysis is extended to include not only the third cross-sectional moment but also the cross-sectional kurtosis and jointly with aggregate idiosyncratic variance.

Also, some authors have found evidence that shows idiosyncratic variance (cross-sectional variance or dispersion) can be seen as a state variable in Merton's (1973) Intertemporal CAPM (ICAPM) framework, as it has shown to be useful in predicting value, size and momentum premia. Guo and Savickas (2006) analyze whether average idiosyncratic volatility helps predict the size, value, and momentum premia and find that it does positively forecast value premium jointly with market volatility. Moreover, Guo and Savickas (2008) find that value-weighted average idiosyncratic variance has a positive and statistically significant predictive power for one-quarter ahead value premium for four of the G7 countries (U.S., Germany, Japan, and UK) controlling for stock market volatility and it acts as a proxy for the volatility of the value premium.

Stivers and Sun (2010) show that cross-sectional dispersion positively (negatively) forecasts the value premium (momentum premium) even when controlling for other macroeconomic variables and Angelidis, Sakkas and Tessaromatis (2015) show that higher cross-sectional dispersion predicts higher value premium, and lower momentum premium and market returns using a twelve-month horizon. Therefore, evaluate if cross-sectional variance has a predictability power of risk factors from an emerging country view may be valuable for a better understanding of what aggregate idiosyncratic risk reflects.

Secondly, the behavior of cross-sectional variance by levels of credit ratings in both stock exchanges is investigated to understand whether the idiosyncratic variance obtained through firms in different categories of credit risk matters to predict market returns. Aditionally, due to limitations in the number of Brazilian firms rated by Standard & Poor's, this study follows Brown and Ferreira (2004) for the U.S. market and Angelidis and Tessaromatis (2008) for the UK market, and disentangles the cross-sectional variance behavior by size, measured by the stock's market capitalization, in the Brazilian case. Angelidis and Tessaromatis (2008) show that the idiosyncratic risk of small stocks is more relevant than those related to all stocks and big stocks only and, hence, this approach can also provide meaningful insights to the association between aggregate idiosyncratic risk and market returns in Brazilian stock market.

Finally, focus shifts to the cross-sectional effects of expected idiosyncratic volatility on expected returns as there is conflicting evidence on whether it has effects on the cross-section of Brazilian stocks returns. Both in Fu (2009) and in Fink, Fink and He (2012), among others, the returns distribution is set to be the gaussian despite the empirical observation that the financial series rejects the normality hypothesis.

Although the assumption that the mean and variance, first and second moments of the return distribution, are sufficient to characterize it underlies central models in Finance, such as Markowitz's portfolio theory (1952) (JONDEAU; ROCKINGER, 2006) and also the normality hypothesis in the CAPM model (PEIRÓ, 1999; VERHOEVEN; MCALEER, 2004), empirically, it is verified that the returns series, in general, do not meet this characteristic and are therefore not normal (VERHOEVEN; MCALEER, 2004; JONDEAU; ROCKINGER, 2006; SU; LEE; CHIU, 2014; FEUNOU; JAHAN-PARVAR; TÉDONGAP, 2014).

Motivated by the conflicting evidence regarding the relation between forecasted idiosyncratic volatility and expected returns, the last attempt in this thesis is to conduct an empirical investigation of expected idiosyncratic risk from the general setting of an EGARCH model assuming a non-gaussian distribution for the innovation process, namely the Skewed Generalized Error Distribution (hereafter Skew-GED) as an effort to overcome the potential issues generated by simply assuming a normal distribution when the literature has empirically shown that significant skewness and kurtosis are exhibited in stock returns.

Overall, the main analysis relies upon the investigation if idiosyncratic risk in both aggregated or firm-level can be related to returns, if there is a positive, negative or non-significant association between them.

1.2 Objectives

1.2.1 General Objective

Investigating the role of aggregate idiosyncratic variance, as proxied by the cross-sectional variance (CSV), in predicting market returns and when disentangling it by the firm's credit ratings in the U.S. and by credit ratings and size in the Brazilian stock market as well as understanding if expected idiosyncratic volatility is related to expected returns.

1.2.2 Specific Objectives

- Analyzing if one-day/one-month-ahead market returns can be predicted by cross-sectional variance in recessions and expansions, up and down markets and subperiods;
- Understanding if investor sentiment can affect the ability of cross-sectional variance in predicting market returns;
- Verifying if cross-sectional variance jointly with cross-sectional skewness and kurtosis predict one-day/one-month ahead market returns;
- Analyzing if cross-sectional variance can predict risk factors omitted from the CAPM model;
- Investigating if cross-sectional variance constructed according to the firm's credit ratings and size are useful in predicting market returns.
- Analyzing abnormal returns of portfolios using forecasted idiosyncratic volatilities obtained at the firm-level;
- Determining whether and how expected returns are associated to expected idiosyncratic volatility using a non-normal return distribution.

1.3 Justification

It is well-known that the risk of a stock is split into two components: the systematic part that is linked to the portion which cannot be removed even if investors try to have a well diversified portfolio, and the unsystematic risk, which represents the one based on the firm-specific risk. Modern Portfolio Theory assumes that investors are rational and diversify away the firm-specific risk. The CAPM model is based on the work of Markowitz (1952, 1959) and its underlying assumptions reflect that investors demand a risk premia according to the exposure to the correlation of the portfolio returns with the market returns, evidenced in the systematic part.

Under-diversification theories, as in Merton (1987), posit that investors do not fully diversify away their exposure to the unsystematic risk, also called idiosyncratic risk and commonly measured by the idiosyncratic volatility, and indicate as one of the reasons for this the existence of transaction costs and limited information. The fact that stocks may have short-sell constraints is also a reason for this phenomenon.

Some empirical evidences presented by Goetzmann and Kumar (2008) sustain that investors bear the idiosyncratic volatility risk as shown by their low degree of diversification. Specifically, the authors use American data during the period between 1991 and 1996 to analyze portfolio positions of 62.387 households who have traded stocks in their database. Their findings support that a 4-stock portfolio is held by an average investor, more than 25% of the portfolios analyzed consisted of 1 stock and more than 50% is concentrated in fewer than 3, even though they comment that a conservative belief for an investor to obtain an adequate level of diversification is at least 10-15 stocks, while Statman (1987) indicates it to be 30.

Campbell *et al.* (2001) point out five reasons to affirm that types of volatility other than the one linked to the market, specifically the idiosyncratic volatility and the industry-level, are important to: i) investors may fail in diversifying or be delimited by corporate policies; ii) investors who try to diversify may have a large number of stocks in the portfolio which seem to be adequated to eliminate the firm-specific risk, but the way the idiosyncratic volatilities of these stocks composing the portfolio relate to each other contributes to the quality of this risk elimination; iii) arbitrageurs dealing with mispricing of individual stocks;

iv) the importance of firm-specific risk in the event studies context; v) the total volatility influences the price of an option.

In this sense, it is important to evaluate the role of idiosyncratic variance itself and when considered by levels of ratings. Avramov *et al.* (2007, 2009, 2013) have shown the ratings contribution in determining the momentum profitability, explaining the negative tradeoff between risk-return and, more recently, its effects on eleven anomalies which are known to generate abornaml returns. Therefore, it may be that idiosyncratic variance of distressed firms, as proxied by their ratings level, can be useful in predicting market returns and this approach is pursued here.

Additionally, a separate strand of literature investigates the role of higher (co)moments (KRAUS; LITZENBERGER, 1976, FANG; LAI, 1997, DITTMAR, 2002) in explaining stock's returns based on non-normality in the asset return distributions (VERHOEVEN; MCALEER, 2004; JONDEAU; ROCKINGER, 2006; SU; LEE; CHIU, 2014; FEUNOU; JAHAN-PARVAR; TÉDONGAP, 2014). Hence, this thesis uses a more flexible distribution to accomodate non-normal distribution (Skew-GED) aiming for a better assessment of the relation between forecasted idiosyncratic volatility and expected returns.

Moreover, the Brazilian stock market is the largest in Latin America with a market capitalization of more than R$ 2,7 trilions (US$ 885 bilions) and the number of retail investors has increased over time, from only 85.249 in 2002 to 569.894 in 2017 and achieving its maximum in 2010 with 610.915. But, when comparing to the U.S. stock market, it can be seen that these numbers are still low. This makes it interesting to evaluate if the aggregate idiosyncratic risk has a different behavior from an emerging market perspective which has produced a few limited evidences regarding this topic whereas the main body of literature in this field is concentrated in the U.S. and some other developed countries.

1.4 Contributions

This thesis adopts the cross-sectional variance (CSV) to understand the behavior and ability of idiosyncratic volatility to predict market returns. CSV is a useful methodology which has the great advatange of being a model-free and easily obtained in any frequency; it was also shown to be highly correlated with other well-used measures of idiosyncratic variance found in the literature, as the ones presented in Goyal and Santa-Clara (2003), Ang *et* *al.* (2006) and Bali, Cakici, and Levy (2008). The contributions presented here start by replicating the CSV as given in Garcia, Mantilla-García, and Martellini (2014) to the United States (U.S.) stock market and extend it to a more recent period, as well by controlling the investor sentiment. Moreover, this work intends to fill a gap in the finance literature by verifying the ability of cross-sectional variance (dispersion) of U.S. rated firms over time in predicting one-day/one-month ahead market returns.

Considering the Brazilian stock market, some of the main findings regarding cross-sectional variance and its correlation with traditional measures are pursued. The analysis of wheter it shows a trend over time is considered, as well as the study of idiosyncratic variance during expansion and contractions periods, up and down markets, and subperiods. Also, its predictive power is investigated in daily and monthly frequencies using equal- and value-weighted weighting schemes, adding the market variance and investor sentiment jointly, and including cross-sectional skewness and kurtosis in the same sense as the former is shown in Garcia, Mantilla-García, and Martellini (2014). The ability of cross-sectional variance in predicting omitted variables from CAPM model is also studied. Additionally, herein is also investigated the cross-sectional variance by the firm's credit rating, but due to the limitation in the number of Brazilian rated firms by Standard & Poors (S&P), the analysis is extended by decomposing it by size.

The study conducted by Sehgal and Garg (2016) was the only paper found which applies the same measure (cross-sectional variance) used here to obtain aggregate idiosyncratic variance in Brazil. But this work differs from theirs as their focus falls on higher cross-sectional moments and the performance of portfolios sorted on these characteristics.

Finally, it is intended to contribute to the literature by investigating whether and how conditional idiosyncratic volatility is associated to expected returns for Brazilian stock without a forward-looking bias, through an EGARCH model specification and using a more flexible distribution to account for non-normality in the return distribution.

1.5 Research Hypotheses

In order to meet the main objective, seven research hypotheses were formulated, as follows:

i) Relevance of aggregate idiosyncratic risk in predicting market returns:

$H1$: CSV plays a relevant role in forecasting one-day/one-month ahead market returns.

ii) Influence of investor sentiment in the relation between aggregate idiosyncratic risk in predicting market returns:

$H2$: Investor sentiment significantly affects the relation between CSV and market returns.

iii) Role of aggregate idiosyncratic risk in predicting market returns in the presence of higher cross-sectional moments

$H3$: CSV is relevant in forecasting one-day/one-month ahead market returns in the presence of cross-sectional skewness and kurtosis.

iv) Forecasting power of aggregate idiosyncratic risk for size, value, and momentum factors, used in the Carhart (1997) model, and also, more recently, given in Fama and French's (2015) five-factors, the investment and profitability premiums.

$H4$: CSV is significant in forecasting one-month ahead risk factors omitted from the CAPM model, such as size, value, momentum, investment and profitability premiums.

v) Usefulness of a measure of aggregate idiosyncratic variance based on credit ratings for the U.S. and Brazilian cases and size only for the Brazilian stock market in predicting market returns.

$H5$: CSV by levels of credit ratings is relevant to forecast one-day/one-month ahead expected market returns.

$H6$: CSV constructed by size (small versus big) is relevant to forecast one-day/one-month ahead expected market returns.

vi) Association between expected idiosyncratic volatility estimated using only the data available to the traders through an out-of-sample approach and with a more flexible non-normal distribution (Skew-GED) for the error term and expected returns using a Brazilian sample of stocks.

$H7$: Expected idiosyncratic volatility at the firm-level is related to expected returns.

1.6 Limitations

The analysis and findings presented hereafter are not free of limitations. There are some authors who do not support using credit ratings as proxy for financial distress of firms, using preferably the Altman Z ratio, where Altman (1968) reports that stocks characterized by lower values of the ratio have higher returns than those with a higher value of the ratio. Due to unavailability of data, this work focused only on firms rated by Standard and Poor’s. Also, the intention here was not to present a formal discussion on idiosyncratic variance trend, therefore a formal investigation about the reasons why it has a trend or not in both stock markets was not pursued here.

Additionally, specifically to the Brazilian market, the timespan covered in the analysis is very short when compared to those provided by WRDS on American data, resulting in a disproportional analysis and, more importantly, which may not reflect the true relationship between the studied variables. Lastly, the proxy used to represent the investor sentiment is limited in terms of comparison as many authors follow Baker and Wurgler’s (2006) investor sentiment index, which also happened with the United States sample.

1.7 Structure of the Thesis

The structure of the remainder of this thesis is as follows. The next section presents the literature review encompassing the traditional market model and multifactor models and explores how the discussion concerning idiosyncratic risk and return relationship has evolved. Section 3 describes the methodological procedures adopted whereas Section 4 presents and discuss the results. Section 5 summarizes the findings and points to possible future research.

2 LITERATURE REVIEW

2.1 Asset Pricing Models

The development of the Finance field is very recent and concentrated mainly in the 1950s. Markowitz's (1952) work provided the basis for the development of the Modern Theory of Finance, which refers to the investor as "rational" and tries to maximize his utility function. The author defines the expected return and the risk associated to the portfolio as follows:

$$E(R_p) = \sum_{i=1}^{n} w_i E(r_i) \qquad \text{Equation 1}$$

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \qquad \text{Equation 2}$$

Where: $E(R_p) =$ expected return of the portfolio *p*; $E(r_i) =$ expected return of asset *i*; $w_i =$ weight allocated to asset *i* in the portfolio *p*; $\sigma_p^2 =$ the variance (risk) of the portfolio *p*; $\sigma_{ij} =$ covariance between the assets *i* and *j*.

Therefore, the portfolio return is obtained by the weighted average of each asset's return by its corresponding weight in the portfolio. However, this formulation is not extended to the risk calculation as the covariance between the pair of any two assets within the portfolio is taken into account and its contribution is essential to an increase or decrease in the level of risk. More specifically, the risk of a portfolio is given by the weighted average of each asset's variance of return by its weight in the portfolio as well as by the covariance between the assets return.

Underlying to Markowitz's (1952, 1959) work is the idea of diversification, which is achieved by investing in assets with different characteristics and responding in a different manner or in a contrary direction to possible sources of risk. When two assets move exactly in the same way, i.e. when they are perfectly correlated, there is no gain from diversification and the risk of a given portfolio simply becomes the weighted average of their return variance by the corresponding weight. Therefore, diversification represents to the investor the minimization of the risk of the portfolio, which is completely obtained when the assets have a perfectly negative correlation. However, there will be a level where the inclusion of assets even with a very low degree of correlation will not contribute to reducing the risk, and, hence, there exists a part of the total risk that cannot be eliminated through diversification.

This part of the risk is the focus of the Capital Asset Pricing Model (CAPM) because it assumes that investors diversify their portfolios and, hence, only systematic risk matters to investors. The CAPM developed by Sharpe (1964), Lintner (1965a), and Mossin (1966), based on the mean-variance approach outlined in Markowitz (1952), is one of the most used asset pricing models academically and by market practioners.

The model that is benchmark in asset pricing (SHIH *et al.*, 2014) associates the expected return and (systematic) risk in an influential and straightforward interpretation (MERTON, 1973; FAMA; FRENCH, 2004), but it is subject to criticism, as will be shown below, though it remains widely used for estimating the cost of capital, performance evaluation (individual assets, investment funds), portfolio diversification, investment valuation, among other purposes (GALAGEDERA, 2007).

The CAPM relates the expected return of a given investment (asset or portfolio) as a linear function of its risk based on Markowitz's (1952, 1959) mean-variance theoretical framework (FAMA; FRENCH, 2004; SHIH *et al.*, 2014). Sharpe (1964) argues that up to that moment there was no specific theory which linked risk to its influences. Therefore, the author asserts that "*lacking such a theory, it is difficult to give any real meaning to the relationship between the price of a single asset and its risk*" (SHARPE, 1964, p. 426). With this, Sharpe (1964) proposes a theory of equilibrium of asset pricing under risk conditions.

According to Fama (1968), the assumptions underlying the CAPM are as follows: i) investors are risk averse and maximize expected utility, as well as being able to make optimal portfolio decisions based on the mean and standard deviation; ii) the investment decision horizon is the same for all investors and there is a mean and a variance of the expected return from this common horizon; iii) the capital market is perfect in the sense that the assets are divisible, there is no transaction costs or any other fees and taxes, the information is disseminated without distinction to all investors and both lending and borrowing rates are identical; iv) homogeneity of possible investment sets and expectations (all investors are assumed to have the same expected mean and variance parameters). Sharpe (1964) also notes that even though these assumptions are not practically verified, the main test of a theory does not lie in the constraints in which the model is built, but rather in how satisfactory its implications are.

Sharpe (1964) extends the mean-variance approach adding the risk-free asset return and the beta, which represents the marginal risk, i.e. the risk which is added to the portfolio by the inclusion of a new asset. Assuming that investors have diversified their investments, variance becomes a non-coherent risk measure as only a portion of the total risk, captured by the beta, matters. This occurs because the total risk can be divided into two parts: non-systemic (diversifiable, idiosyncratic) risk, which can be eliminated through diversification, and systemic (non-diversifiable) risk, which cannot be eliminated and is common to all investors, for instance changes in the risk-free rate, rate of inflation, exchange rate, among many others. The beta is expected to represent the latter. Therefore, the CAPM equation is as follows:

$$r_{it} - r_{ft} = \alpha + \beta\left(r_{mt} - r_{ft}\right) + \varepsilon_t \qquad \text{Equation 3}$$

Where: r_{it} refers to the asset's return in period t; r_{ft} represents the risk-free asset return in period t; r_{mt} is the market return in period t; ε_t is the model's residual.

Therefore, the return on the asset is given as a function of the return on the risk-free return plus a market risk premium (the additional to the portfolio return on the return of the risk-free) multiplied by the asset's beta, β. In addition, the expected value of the error term is zero. The higher the beta is, the greater the expected return.

The market portfolio's beta is equal to 1. All other betas are characterized as less risky or riskier compared to it. In particular, $\beta < 1$ represents that a given asset has lower risk than that of the market portfolio; $\beta > 1$ means that a given asset has higher risk than that of market portfolio. Assuming that the investor's portfolio is fully diversified, the error term in the CAPM equation would play no role, indeed this is the model's assumption. To portray this more clearly, when put together the systematic and non-systematic parts, then the total risk is:

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma^2\left(\varepsilon_i\right) \qquad \text{Equation 4}$$

Where: σ_i^2 = variance (total risk) of asset *i*; $\beta_i^2\sigma_m^2$ = systemic risk; $\sigma^2(\varepsilon_i)$ = firm-specific risk or idiosyncratic risk.

According to Fama and French (1992, p. 427), the central point of the model lies in the fact that the market portfolio is mean-variance efficient according to the concept proposed by Markowitz (1959). In addition, this is indeed a single factor model because the expected return is directly related to a single risk measure estimated by the beta.

However, the CAPM has been criticized in the literature since its publication. Galagedera (2007, p. 822) mentions that "*a growing number of studies found that the cross-asset variation in expected returns could not be explained by the systematic risk alone*". Fama and French (2004) suggest that the poor performance of the CAPM may be relative to its simplified assumptions or the difficulties of empirically testing it. Black (1972), for instance, points out that the most restrictive assumption is that investors can take any long or short position. However, the author shows that even in the absence of the risk-free asset or by imposing restrictions in the positions assumed in it and unrestricted positions in risky assets, the CAPM remains as a model.

Roll (1977) states that the model should be seen more as a market efficiency test than an asset pricing model, since the author shows that testing the CAPM has only the purpose of analyzing whether the market portfolio is mean-variance efficient and the impossibility of accessing this portfolio, since it should comprise all assets and human capital; i.e. there is no way to observe it empirically.

In a different context, Ross (1976) seeks to develop an arbitrage-based asset pricing model, the Arbitrage Pricing Theory (APT) which, unlike CAPM, was not formulated as a model of equilibrium pricing. Since the goal is to create a model capable of explaining the variation of returns, the APT suggests that there are systemic factors capable of capturing such relationship without specifying them.

The CAPM states that there is a linear relationship between the return of the asset with a single factor (the excess of the market portfolio), measured by its beta, whereas the APT model assumes there are other factors that linearly explain the return of the asset. In addition, there are differences in the assumptions of the models as the former is based on the market equilibrium and the second in the arbitrage theory.

Another strand of the literature seeks to find out relevant variables to explain the return variation lead to the development of several factorial models based on empirical evidences. Banz (1981) argues that the CAPM is a poorly specified model by analyzing that smaller companies earn higher returns than big companies, denominating this empiric observation as the size effect. Chan, Hamao, and Lakonishok (1991) examine four fundamentalist variables to determine whether they could be linked to japanese stock returns, namely the earnings yield, size, book-to-market and cash flow yield.

The authors identified that the book-to-market and cash flow yield are the most significant variables to explain the returns. The tests conducted by Chan, Hamao, and Lakonishok (1991) also confirm the existence of the size effect in the Japanese stock market, although sensitive to the model specification. Rosenberg, Reid and Lanstein (1985) identify the book-to-market variable as significant and positive in explaining returns.

Fama and French (1992) also conclude that the CAPM's beta failed to capture the returns variability in a satisfactory way. The authors corroborated the arguments raised by Stattman (1980), Basu (1983), and Bhandari (1988), who identified other significant factors in the explanation of returns. Accordingly, they presented a multifactorial model. Fama and French (1993) use the CAPM's beta related to the risk and add the size and value premium in their newly proposed model due to their ability to explain the returns.

Specifically, the size premium is captured by the *Small minus Big* (SMB) factor, which was termed after the observation by the authors that small stocks have higher returns than big stocks. Also, the value premium is given by the *High minus Low* (HML) factor, which in turn is built in the idea that stocks with high book-to-market ratio have higher returns than those with lower book-to-market. This model is named as Fama and French (1993) three factor model (hereafter FF-3).

Overall, the authors find that the FF-3 is preferable to the CAPM model. Nevertheless, Carhart (1997) extends the FF-3 specification to include the *Momentum* (MOM) factor. This factor is defined as a strategy that longs stocks with good performance in the last few months (short term) and shorts those with bad performance in the analyzed period. This strategy is found to be profitable by Jegadeesh and Titman (1993).

Therefore, it represents the idea that stocks with higher returns will continue the favorable trend in the short term and those that lower returns will continue in the downtrend. Carhart (1997) indicates that with the inclusion of momentum factor, the new model significantly reduces the average pricing errors when compared to the CAPM and the FF-3 as well as explains almost the pricing error behavior. This model is known as the Carhart (1997) model (hereafter FFC).

More recently, Fama and French (2015) developed a new asset pricing model known as the Fama and French (2015) five-factor model (hereafter FF-5) which adds to the FF-3 two additional factor-mimicking portfolios based on investment and profitability. Titman, Wei, and Xie (2004) document a negative relation between abnormal capital investments and returns. Moreover, Novy-Marx (2013), using the ratio of a firm's gross profits-to-assets as a mesaure of profifability, argues in favor of its importance showing that it has pratically the same efficiency as the book-to-market in explaining the cross-section of average returns.

In this sense, Fama and French (2015) argues that the choice for these new factors is based on the dividend discount model and add both investment and profitability to the market, size, and value factors in their FF-5 model. More specifcally, the profitability premium, based on the *Robust minus Weak (*RMW*)* factor, represents that stocks with high profitability earn higher returns. In contrast, the investment premium, based on the *Conservative minus Aggressive* (CMA) factor, reflects the observation that low investment stocks earn higher returns than those with high investment. It is noteworthy to mention that the HML factor loses explanatory power and becomes redundant when these two new factors are accounted for, which can be related to their sample (FAMA; FRENCH, 2015). Also, the authors report the model's inability to explain small stocks with returns that perform like those of firms with high investments (aggressive), though having low profitability (weak). Overall, the authors point out that FF-5 outperforms the previous FF-3 model.

2.2 Aggregate Idiosyncratic Volatility or Variance

The study of idiosyncratic risk follows the development of the CAPM model, which assumes that investors require a premium for bearing only the systematic part of the total risk, measured by the beta. Lintner (1965b) and Douglas (1969) argue that idiosyncratic risk matters as a source of risk, while Miller and Scholes (1972) report problematic methodological issues affecting their results, and Fama and MacBeth (1973) point out that the standard deviation of residuals from the market model (idiosyncratic volatility) does not play a significant role in explaining average returns in their cross-sectional estimate testing the

CAPM model. In this sense, Lehmann (1990) provides several econometric specifications and shows that in the full sample period the idiosyncratic risk is (positively) significant.

On the other hand, Campbell *et al.* (2001) study the decomposition of volatility into the market, industry, and firm-level without the need of estimation of the coavariance structure or betas during the period (July 1962 to December 1997). The authors show a rising trend in idiosyncratic volatility, whereas market and industry volatilities show stable behaviour during the period and that aggregate idiosyncratic volatility is useful in forecasting economic activity. The declining correlation between stocks evidenced by the authors points to about 50 stocks being nescessary for a diversified portfolio. In accordance to this finding, Kearney and Poti (2008) affirm that 166 stocks are needed to reduce idiosyncratic risk to a 5% level in a portfolio in 2003, while this number was 35 in 1974, using european stocks.

Xu and Malkiel (2003) confirm with U.S. data a positive trend upwards in the idiosyncratic volatility during the period from 1952 to 1998 and postulate that both the number of institutional ownership and expected earnings growth are positively related to idiosyncratic volatility. In other words, the relative number of institutional ownership in the market has increased significantly over time and affects positively the firm-specific risk and, additionally, idioynscratic volatility is higher for firms whith higher perspectives of growth, which renders them riskier because of their dynamic nature in a competitive environment.

Meanwhile, Brandt *et al.* (2010) reveal that the increase in idiosyncratic volatility identified by Campbell *et al.* (2001) must not be seen as a time trend, but rather as an episodic phenomenon as the evidence presented in their papers show that by 2003 idiosyncratic volatility had already decreased to its level before 1990. The authors link this idiosyncratic volatility behavior to low-priced stocks with high retail trading[1]. Pástor and Veronesi (2003) indicate that the idiosyncratic volatility of profitability and uncertainty about profitability contributed to the behavior of the aggregate idiosyncratic risk illustrated by Campbell *et al.* (2001) and, hence, are aspects which should be considered to explain their results.

In contrast, Bekaert, Hodrick, and Zhang (2012) point out that there is no upward trend in idiosyncratic volatility while studying a U.S sample of stocks using an extended

[1] It is noteworthy that Brandt *et al.*'s (2010) explanation, as already highlighted by the authors, about the role of retail investors in the understanding of the increase in aggregate idiosyncratic risk contrast with Xu and Malkiel's (2003) results, which relate it to the amount of institutional investors. To a deeper understanding, Brandt *et al.* (2010) replicate Xu and Malkiel's (2003) estimations and additionally split the sample into low-priced stocks and high-priced stocks. Their results imply a negative association between idiosyncratic volatility and institutional ownership when looking at the sample of low-priced stocks.

period (from 1964 to 2008) and the same pattern is observable in more 23 developed markets starting on 1980. Indeed, there is a rise in the average idiosyncratic volatility when focusing on 1997, as done in Campbell *et al.* (2001), but Bekaert, Hodrick, and Zhang (2012) provide evidence that this is not the case when expanding the analysis to cover a longer sample. Furthermore, their findings show that idiosyncratic volatility across the analyzed countries is highly correlated.

There is also conflicting evidence regarding the idiosyncratic risk importance. Goyal and Santa-Clara (2003) contribute to the debate affirming that aggregate risk (mostly driven by idiosyncratic variance) matters during the period from 1962 to 1999. Wei and Zhang (2005) revisit Goyal and Santa-Clara's (2003) finding through additional data covering the period up to December 2002. The authors compute their average volatility and cross-sectional variance (which is also an estimate of idiosyncratic risk) measures. One of the main intriguing point refers to the association between the value-weighted returns and the lagged equally-weighted average volatility which is shown to have a positive and statistical significance in Goyal and Santa-Clara's (2003) study.

However, Wei and Zhang (2005) do not find it to be robust accross different subsamples and when considering the extended period. As opposed to the focus on the relation between value-weighted average return and equally-weighted average volatility, Wei and Zhang (2005) report that the equally-weighted average return and the equally-weighted average volatility or cross-sectional variance have a positive and significant relation, though that does not generate economic gains either.

Angelidis and Tessaromatis (2008) examine the forecasting power of lagged idiosyncratic volatility in predicting equal- and value-weighted market returns and the return on SMB and HML portfolios, following Fama- French (1993) with United Kingdom data. The intuition behind the ability of idiosyncratic risk in forecasting these factors lies in the Intertemporal CAPM (ICAPM) model of Merton (1973), which does not discriminate the most relevant state variables to account for as a hedge portfolio.

Moreover, to estimate the idiosyncratic risk, the difference between the aggregate variance measure used by Goyal and Santa-Clara (2003) and the variance of the market is obtained and, similar to Brown and Ferreira (2004), the authors split idiosyncratic variance into large and small components based on the market capitalization of stocks. In this sense, three measures of aggregate idiosyncratic risk are employed: one based on all stocks, a second one containing stocks with 80% percent of the total market capitalization and, finally, a third one consisting of stocks with the 20% smallest stocks in market capitalization.

An overview of the author's results evidences that the idiosyncratic volatility of small stocks play a more important role than those of larger stocks. When the value-weighted market return is regressed on lagged measures of market volatility and alternative measures of idiosyncratic variance, the authors cannot find any significant relation; however, using the equally weighted market return the idiosyncratic variance based on small stocks and all stocks reveal a positive significance at the 10% level, a marginal predictive importance.

The results concerning the forecasting power of the SMB portfolio returns by both value-weighted idiosyncratic variances of small and large stocks in the model specification show that the idiosyncratic measures are significant, but only that of the small stocks is positive and significant at the 5% level, whereas no measure seems to be adequate in forecasting the HML portfolio returns as they are undistinguishable from zero.

Angelidis and Tessaromatis (2008) findings do support that value-weighted idiosyncratic variance of small and large stocks are more relevant to predict SMB portfolio returns than other state variables as dividend yield, default premium, detrended risk-free rate and also expected and unexpected illiquidity. Furthermore, when validating the robustness of their results across subsamples, it can be seen that the corrected Newey-West (1987) t-statistics of small stocks idiosyncratic volatility show itself positive and significant specially after 1994, while, for larger stocks, no statistical significance accross the years is shown. Therefore, their findings shed light on the importance of aggregate idiosyncratic variance, but only that generated from small stocks.

Jiang and Lee (2006) examine the time-series effect between idiosyncratic volatility (shocks) and market returns. Their work suggests a positive relationship even after controlling for some well-known state variables seeking to take into account changes in the investment set, as motivated by the ICAPM of Merton (1973) and following Goyal and Santa-Clara (2003) and Bali *et al.* (2005).

The authors use both equal- and value-weighted schemes as well different measures of volatility: market volatility, aggregate idiosyncratic risk measure (as proposed in Goyal and Santa-Clara, 2003), which is mostly idiosyncratic, a measure of idiosyncratic variance according to a constant mean return model, and also from the variance of residuals regression based on the CAPM, Fama- French (1993), and Carhart (1997) models. The central point in their estimation is the inclusion of orthogonalized innovations together with lagged idiosyncratic variance, which is done due to the high persistence detected in the latter and, hence, making inadequate the use of its own lagged values alone. Thus, Jiang and Lee (2006) try to correct for autocorrelations in the firm-specific risk.

Guo and Savickas (2006) highlight that the reason for some previous work having failed in finding a posivite tradeoff between risk and return is due to a omitted variable problem. When market volatility and idiossyncratic risk are both taken in the analysis, the authors conclude that both are significant, though with opposite signals – the former is positive and the latter is negative – and also report a correlation of 0.77 between them using quarterly data.

Guo and Savickas (2006) briefly verify that aggregate idiosyncratic risk has an upward trend, as discussed in Campbell *et al.* (2001), and include a detrended version of idiosyncratic risk or a linear time trend variable to take this characteristic into account in their time-series analysis. The main measure to estimate idiosyncratic volatility is obtained through the residuals of a regression relative to the Fama and French (1993) three-factor model applied in the Goyal's and Santa-Clara's (2003) approach (raw returns are replaced by the already computed residuals) and a market capitalization weighting scheme.

The authors' findings show that aggregate idiosyncratic variance has forecasting power for market returns both in-sample and out-of-sample and a negative and statiscal significant association in forecasting each one of the ten excess portfolios returns built on the previous period's idiosyncratic variance. Guo and Savickas (2006) also seek to forecast the return on SMB, HML, and momentum factors and find that aggregate idiosyncratic volatility does forecast HML portoflio return jointly with market volatility.

Two different, important approaches conducted by Guo and Savickas (2006) are the study of the forecasting power of excess market return along with the consumption-wealth ratio (CAY) of Lettau and Ludvigson (2001) and distinct aggregate liquidity measures, whereas these variables are found to play an important role in the asset pricing literature. When using CAY, their findings point that idiosyncratic variance remains negative and siginificant while the t-statistics of both variables decrease, however, when considering its detrended version or the variable itself with the inclusion of a linear time trend, then it becomes insignificant while CAY and market variance retain their predictive power. The authors conclude that it is due to the negative association of CAY and idiosyncratic variance that the latter helps to forecast market returns. In addition, the use of alternative measures of aggregate liquidity also shows that idiosyncratic variance is still significant in different model specifications while they seem to be linked.

Overall, Guo and Savickas (2006) consider the negative relationship between market returns and idiosyncratic variance to be in accordance with the hypothesis that it reflects a problem due to the non-inclusion of the liquidity variable, which Guo (2004) posits to have a possible negative sign, or with the hypothesis that idiosyncratic variance is a proxy for the dispersion of opinion, in which the authors refer to Cao, Wang, and Zhang's (2005) findings, in which it has a positive relation to the market volatility and a negative one to the excess market return.

2.3 Firm-level Idiosyncratic Volatility

Aware of the idiosyncratic risk which may be influencing the investor portfolio, Merton (1987) and Malkiel and Xu (2002) suggest that a higher return should be expected, i.e. a positive relationship between idiosyncratic volatility and future returns, in contrast to the standard CAPM model in which idiosyncratic risk plalys no role. Barberis and Huang (2001) also argue that higher idiosyncratic risk should command a higher return. In early work, Lintner (1965b) argues in favor of the relevance of the variance of residuals in the cross-section estimations and Lehmann (1990) also provides support of its positive relation with returns using the entire period. But this topic has been extensively debated in the asset pricing field given the recent conflicting result with what the theory suggests.

In this sense, the findings presented by Ang *et al.* (2006, 2009) show that realized idiosyncratic volatility is negatively associated with expected returns, called the "idiosyncratic volatility puzzle". Ang *et al.'s* (2006) findings show controlling some of the commonly cited variables as relevant in the asset pricing field, such as size, book-to-market, momentum, liquidity and dispersion of analysts' forecast were not able to dismiss the puzzle.

The authors use the standard deviation of residuals of the Fama and French (1993) three factor model to measure idiosyncratic volatility and group stocks into quintles based on their previous month's innovation, resulting in a significant raw monthly return difference of -1.06% (and an alpha relative to the Fama and French three factor model equal to -1.31%)

between the group of stocks in the highest quintile (highest level of idiosyncratic volatility) and the group in the lowest quintile (lowest level of idiosyncratic volatility).

Despite the controversial result using U.S. data, Ang *et al.* (2009) demonstrate that it is not a unique anomaly found in American stock exchanges. The authors extend their research to include 23 developed countries around the world and confirm the uncovered anomaly in each of the G7, as well in other stock exchanges analyzed. Moreover, Ang *et al.* (2009) point out that, due to the observed common aspects regarding the spread of portfolios ranked in the highest to the lowest idiosyncratic volatility, that there is an underlying mechanism to this anomaly that cannot be easily removed. The most important point is that if idiosyncratic risk matters, then traditional models such as the CAPM or the Fama and French three factor model are not enough to describe the stock returns.

Notwithstanding, several other studies have investigated the role of idiosyncratic volatility and have found either a positive or non-significant association with average returns. Malkiel and Xu (1997) show a positive relation between idiosyncratic volatility and average returns by sorting stocks into decile portfolios, though the authors do not provide the statistical significance of their results. Malkiel and Xu (2002) revisit the Fama and MacBeth (1973) methodology, which gives support to the non-significance of idiosyncratic volatility as a risk factor, but the authors found, using an updated version of CRSP files and increasing the number of portfolios, that there is some evidence suggesting a positive and significant firm-specific risk in the cross-sectional regressions; the Fama and French (1992) approach was also implemented in their paper, showing the idiosyncratic risk importance.

Bali and Cakici (2008) argue that there is no significant relation between the expected returns and idiosyncratic volatility as it becomes evident from their analysis that the sensibility of the results to different weighting schemes, breakpoints used to sort stocks, frequency of the data, and the role of size, price, and liquidity of the stocks. The authors use two different measures to obtain idiosyncratic volatility, specifically one from daily returns using the within-month observations and a second approach from monthly returns using the past 24 to 60 observations to estimate the standard deviation of the residuals relative to the CAPM and Fama-French (1993) model.

Their results suggest alternative directions in value-weighted portfolios regarding the relationship between expected returns and how idiosyncratic volatility is calculated: a negative and significant association using the former and a flat relationship using the latter. A deeper investigation was conducted on the characteristics of both measures and because of the monthly measure was found to be more accurated, Bali and Cakici's (2008) results do not corroborate with Ang *et al.*'s (2006).

Fu (2009) stresses that idiosyncratic volatility, as measured by Ang *et al.* (2006), does not follow a randow walk, but is time-varying. Fu (2009) posits that the correct variable should be the expected idiosyncratic risk estimated based on the same moment in time as the expected returns; turns out that lagged variable does not appear to be a good proxy for expected idiosyncratic volatility.

To this purpose, the author obtains in-sample forecasting of idiosyncratic volatility at the firm-level through the Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH) model of Nelson (1991) with the Fama-French (1993) three factor model in the mean equation and a normal distribution for the error term. By doing so, Fu's (2009) findings suggest a positive and significant relation between expected idiosyncratic volatility and expected returns, in which the spread between the portfolio returns in the highest and lowest idiosyncratic volatility deciles delivers a monthly return of 1.75%. In addition, the author shows that the return reversal of stocks with high idiosyncratic volatility, which are mostly small, helps to solve the puzzle.

Chua, Goh, and Zhang (2010) argue that the relation of unexpected returns (UE) with unexpected idiosyncratic volatility (UIV) may be influencing the estimation of the real relationship between expected returns and expected idiosyncratic volatility, which represents the focus of their research. Hence, this can be a reason underlying the mixed evidence found in the literature. In this sense, the idiosyncratic volatility is computed as Ang *et al.* (2006) and the one-month ahead idiosyncratic volatility is predicted using the coefficients estimated through an Autorregressive model - AR(2), which corresponds to the expected idiosyncratic volatility and, in addtion, the one-month ahead residuals correspond to the UIV.

Indeed, by decomposing the volatility, the authors find a positive correlation between UE and UIV. When taking into account the UIV variable as a control variable for the UE in the regressions, both UIV and expected idiosyncratic volatility are positive and statistically significant in explaining the cross-section of returns even after controlling for size, book-to-market, accumulated returns and liquidity. The result is in accordance with Fu (2009), who also detect a statistically positive relation between expected returns and expected idiosyncratic volatility.

In the same line of work, Spiegel and Wang (2005) utilize the EGARCH with the Fama and French (1993) three factor model in the mean equation, using monthly data to obtain conditional expected idiosyncratic volatility. The authors disentagle the behavior of idioyncratic risk and liquidity confirming their negative interaction in the stock return, though the firm-specific risk effect shows itself to be stronger in the estimated regressions.

Brockman, Schutte, and Yu (2009) apply the same methodology outlined by Fu (2009) in a large sample consisting of 44 countries around the world, including some emergents, such as Brazil. Their Fama-MacBeth (1973) cross-sectional regressions give support to a positive and significant price of conditional firm-specific volatilities in 36 countries and the remaining, with the exception of one country (Luxembourg), are shown to be positive, even though not significant.

Huang *et al.* (2010) attribute the negative effect of realized idiosyncratic volatility in expected returns related to return reversals. But Peterson and Smedema (2011) argue that the negative relation still exists in non-January months even after return reversals are accounted for. Their paper documents the role mispricing plays in Ang *et al.*'s (2006) results, but at the same time the expected idiosyncratic volatility remains positive and significant in both January and non-January months and is more consistent with the idea that it reflects the required investor's compensation for bearing this risk.

Boyer, Mitton, and Vorkink (2010) indicate the idiosyncratic skewness and Han and Kumar (2013) suggest the retail trading proportion. For example, Bali, Cakici, and Whitelaw (2011), based on the evidences that investors poorly diversify their portfolios and also prefer lottery-like assets, which means assets with small probability of a large payoff, point out to the relevance of the maximum return in the previous month (MAX) to account for the idiosyncratic volatility puzzle. The authors show that when both lagged idiosyncratic volatility and MAX are put together, the former presents a positive sign, although not significant in all estimatios, resulting in the impossibility of considering idiosyncratic risk a proxy for the maximum return.

Overall, the evidences presented by Fu (2009) through the measure of conditional firm-specific volatility in opposition to the realized show that the former type of risk exists and commands a positive and significant risk premia, which is in accordance with the theory predicting that due to under-diversification the idiosyncratic volatility should be positively priced.

But there is some disagreement regarding this implementation. Guo, Kassa, and Ferguson (2014) claim that Fu’s (2009) results suffer from a look-ahead bias given the incorporation of one data point (one forward-looking observation), i.e. the author uses data up to time *t* to forecast the idiosyncratic volatility in time *t*. Thus, Guo, Kassa and Ferguson (2014) assert that there is no relationship between conditional idiosyncratic volatility and expected returns when this recent observation is ignored in a truly out-of-sample experiment.

However, Fu (2010), in response to this new evidecence, argues that Guo, Kassa, and Ferguson’s (2014) implementation has two major drawbacks: first, the inclusion of forecasted idiosyncratic volatilities, which does not necessarily have a status reporting that the model in fact achieved convergence; second, in their settings, the maximum number of iterations is 100, while Fu (2010) considers that very restrictive and mentions that 500 iterations are enough. Therefore, the author reestimates the model excluding the most recent observation and accounting for these critiques and indicates that the first positive association remains.

Trying to shed some light on the topic, Fink, Fink, and He (2012) evaluate alternative models that have already been used in a setting which includes realized and expected idiosyncratic volatility measures, as the one proposed by Ang *et al.* (2006) and Fu (2009), and with an alternative information set, as follows: forecasted values from the full dataset (as in Brockman, Schutte, and Yu, 2009); in-sample estimates they mention to be “up to time t to forecast time t” (as employed by Fu, 2009); and out of sample estimates in which the most recent observation is set to be missing (Guo, Kassa, and Ferguson, 2014).

The authors impose a maximum number of iterations equal to 2000 to avoid any problematic issue as the one detected in the Guo, Kassa, and Ferguson’s (2014) procedure. The inferences presented by the authors refute the positve relationship between idiosyncratic volatility and expected returns when conditioning the information set until t-1, i.e. in the out of sample version of the EGARCH measure, while a positve one remains if forward-looking bias is introduced through the use of the most recent observation.

In another strand of the literature, empirical evidences of non-normality in the financial data have appeared, in opposition to the generally assumed normal distribution for the error terms. Wan and Xiao (2014) criticize the use of a normal distribution across all stocks and affirm that, in the presence of non-gaussianity, the EGARCH method of obtaining idiosyncratic volatility in a normal distribution assumption setting may contain finite sample bias. The hypothesis concerning the normality of return innovations, skewness equal to zero and excess of kurtosis equal to zero are rejected for 90.3%, 80.7%, and 87.5%, of the stocks traded on the NYSE, NASDAQ, and AMEX respectively, at the 5% of significance.

Given these evidences, the authors propose a new approach to estimating conditional firm-specific volatility which makes use of quantile regression to accomodate the returns characteristics and does not require any specification for the innovation process. The authors report that relaxing the normality assumption makes it no longer possible to recover the positive association between expected returns and conditional idiosyncratic volatility. Indeed, the monthly average differential return between the stocks with the lowest and the highest conditional idiosyncratic volatility is documented to be 1.53%, in contrast with earlier evidences (Fu, 2009).

Nevertheless, Wan and Xiao (2014) show through the cross-section regressions that controlling the nonlinearity in conditional idiosyncratic volatility (idiosyncratic variance) and conditional skewness in the relationship between expected returns and conditional volatility is found to be statistically positive in stocks which account for more than 99% of the market capitalization. Thus, both variables contribute in the explanation for the puzzle.

In this sense, the results reported here indicate that still no consensus exists on the trend in aggregate idiosyncratic volatility or on its relationship with expected returns. Table 1 briefly presents some of the main findings regarding the study of aggregate and firm-level idiosyncratic volatility in Brazil.

Table 1 – Empirical Evidences on Idiosyncratic Risk

Authors	Period	Idiosyncratic Volatility Definition	Main Results
Campbell et al. (2001)	07/1962 – 12/1997	Indirect method of decomposition	The authors decompose stock volatility at the firm, industry and market level and document that the volatility at the industry and market levels have been stable over time in opposition to the upward trend verified in the average idiosyncratic firm volatility.
Goyal and Santa – Clara (2003)	08/1963 – 12/2009	Total variance	The authors find that the equally-weighted total variance, mostly idiosyncratic, helps to forecast value-weighted market returns with a significant and positive signal.
Guo and Savickas (2006)	1963:Q4 – 2002:Q4	Goyal and Santa-Clara (2003) – Total variance	When both market volatility and idiosyncratic risk are taken in the analysis fo forecast value-weighted market returns to the United States (U.S), the authors obtain that both are significant, though with opposite signals, in which the former is positive and the latter is negative. Guo and Savickas (2008) extend the analysis to the G7 countries and find that adding the contry-specific regressors with those of the United States, with exception of Japan, the negative relation is found to be statistically or marginally significant for the remaining countries.

Continued

Ang et al. (2006)	07/1963 – 12/2000	FF – 3 residuals	The authors show that that realized idiosyncratic volatility is negatively associated with expected returns. Ang et al. (2009) extend their research to include 23 develop countries around the world and confirm the uncovered anomaly in each of the G7 countries, as well in other stock exchanges analyzed.
Galdi and Securato (2007)	01/1999-03/2006	Goyal and Santa-Clara (2003) – Total variance	No relation between realized total (mostly idiosyncratic) variance and portfolio return against the evidences presented by Goyal and Santa-Clara to U.S. stock market showing that equal-weighted total variance predicts value-weighted market returns.
Brockman, Schutte and Yu (2009)	07/1980 – 10/2007	Fu (2009) – EGARCH and FF - 3 residuals	The authors cover 44 countries in their analysis including the Brazilian stock market. Specifically, a positive and statistically significant relation between conditional idiosyncratic volatility and expected returns is recovered, although using realized idiosyncratic volatility there is a negative and insignificant association, which contradicts the evidences presented in Ang et al. (2006) for U.S. where it is negative and statistically significant.

Continued

Angelidis (2010)	12/1994 – 05/2007	Ang et al. (2006) – CAPM residuals	The author studies idiosyncratic volatility across 24 emerging countries including the Brazilian stock market. Angelidis (2010) finds that firm-specific risk in Brazil decreases over the sample period although statistically insignificant. In relation to the predictive ability of idiosyncratic volatility, the author also documents that it is negative and statistically significant while market variance is positive and non-significant.
Mendonça et al. (2012)	07/2005-12/2010	Fu (2009) – EGARCH and FF - 3 residuals	Positive relation between idiosyncratic volatility and returns, though the realized idiosyncratic measure is more relevant than its counterpart. Fu (2009) finds a positive relation between expected idiosyncratic volatility and expected returns and when adding realized idiosyncratic volatility to the model it is still significant though reduces the average slope (-0.02) casting doubts on its profitability.
Costa, Mazzeu and Costa Junior (2016)	01/1996 – 12/2010	Campbell et al. (2001) – Indirect method of decomposition	Negative trend in the firm-level idiosyncratic volatility from september 1998 while the market and industry volatility levels remain stable over the entire time span. This is in opposition to Campbell et al. (2001) findings to U.S. stock market which shows an upward trend in firm-specific risk whereas industry and market volatilities are stable over time.

Note: For this thesis, except the approach followed by Campbell et al. (2001); the remaining were replicated in the aggregate idiosyncratic risk part or at the firm-level analysis and when needed adjusted to the purpose of this work. Further details about the implementation of these variables are found in the methodology section.
Source: Author (2017)

3 METHODOLOGY

3.1 Research Characterization

The present research is characterized as quantitative and classified as descriptive because it aims to describe the possible relation between idiosyncratic risk and returns at the market and firm level, and *ex post facto*, since it makes use of historical data. Therefore, there is no way to change the observed return. According to Creswell (2007), post-positivism claims are used in the quantitative technique, and as the researcher intends to test or verify theories or hypotheses, to make use of non-biased standards, and to measure the numerical information, this is the method employed here.

3.2 Data Description

In order to analyze aggregate idiosyncratic risk in the United States stock market, only daily stock returns of common equity shares (share code 10 and 11) traded on NYSE, AMEX and Nasdaq and their market capitalization acquired from the Center for Research in Security Prices (CRSP) were taken into consideration. Stock returns adjusted for stock splits and dividends were collected from the daily returns file and market capitalization was calculated by multiplying the stock price (PRC) by the number of shares outstanding (SHROUT) available in the database.

The analysis covered the period from July 1963 to December 2014 in the U.S stock market. Equal- and value-weighted CRSP indexes were also collected. The risk-free rate was obtained from the Kenneth French's website[2] as well the risk factors in the Carhart (1997) and Fama and French's (2015) five-factor model.

The main investigation upon the level of credit rating ranged from January 1987 to December 2014 using the United States firms. The CRSP/Compustat Merged Database (CCM) provided by Wharton Research Data Services (WRDS) was used to link the firm's ratings to their corresponding stocks data. The Long-Term Issuer Credit Ratings provided by Standard & Poor's available in Compustat on WRDS was used as the credit ratings variable.

[2] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

The investor sentiment index for American data was collected from Professor Jeffrey Wurgler's website[3] with time series available only from July 1965, therefore the analysis which includes this variable started with this period.

Daily and monthly stock returns of Brazilian common and preferred stocks as well as all variables related to the firms (number of shares outstanding, number of shares traded, book value, operational profitability, trading volume, among others) described here with Brazilian data were obtained from the Quantum® database. The exception is the credit ratings, which come from Bloomberg® database[4] from the Center for Research in Finance of CEFET/MG. Daily stock returns adjusted for stock splits and dividends were employed in the aggregate idiosyncratic risk analysis.

The analysis at the aggregate level of idiosyncratic variance covered the period from January 2000 to June 2016, while, at the individual level, it began in July 1999 because of availability of data to construct the factors. Ibovespa was used to construct the market risk premium in the risk factor constructions outlined in Section 3.3.2 whereas the Interbank Deposit Certificate (Certificado de Depósito Interbancário, CDI) was chosen as a proxy to the risk-free rate.

In the Brazilian case, there is no publicly available investor sentiment index as given in Baker and Wurgler (2006), therefore another proxy was used according to the literature, which is based on investor confidence indexes as Lemmon and Portniaguina (2006) and Schmeling (2009). To this end, monthly time series data of Consumer Confidence Index (Índice de Confiança do Consumidor, ICC) was downloaded from IPEADATA[5]. Moreover, to accomplish this goal, the Index of economic activity (ICB-Br), and inflation rate available from Central Bank of Brazil[6] were also collected to control for macroeconomic conditions which could affect the relation. The data for IBC-Br starts only in January 2003, then the analysis for this part start in this period.

Also, aggregate dividend yield was collected from the Brazilian Center for Research in Financial Economics (NEFIN) of the University of São Paulo[7], and corresponded to the total dividend payments during the last 12 months divided by the market value of equity. The

[3] Available from: http://people.stern.nyu.edu/jwurgler/.
[4] We would like to thank the Center for Research in Finance of Cefet/MG for providing us access to the Bloomberg® database.
[5] Availabe from: http://www.ipeadata.gov.br/Default.aspx.
[6] Available from: https://www.bcb.gov.br/pt-br/#!/home.
[7] Available from: http://nefin.com.br/.

variable is available at the weekly frequency, and then the last observation per month was used to obtain a monthly series. Moreover, it is only available from February 2001, then where it is included the analysis start at that period.

3.3 Analysis Procedure

3.3.1 Sample

The analysis procedure considering the aggregate idiosyncratic variance and using the United States (U.S.) stocks was primarily focused on construction of an aggregate idiosyncratic variance measure as detailed in Garcia, Mantilla-García, and Martellini (2014) and described here in Section 3.4.1, named cross-sectional variance, henceforth, CSV. This measure was chosen as it has two main advantages: not relying in any specific model and being easily calculated for any frequency of data.

Garcia, Mantilla-García, and Martellini (2014) show that CSV is highly correlated with some of the well-known measures of aggregate idiosyncratic variance, such as the one used by Ang *et al.* (2006), Goyal and Santa-Clara (2003), and Bali, Cakici, and Levy (2008), therefore the cross-sectional variance (dispersion) of returns serves as a proxy to estimate the non-observed idiosyncratic variance. Firstly, the procedure to obtain the final sample in both the United States and Brazilian stock market was the same. Each month stocks with missing returns and with non-positive market capitalization at the beginning of the month were dropped.

Secondly, the CSV was constructed according to the firm's credit ratings to investigate if it can help to predict one-day/month ahead market returns. The credit ratings are expressed opinions about the condition and disposition that the obligor has to meet its financial obligations (STANDARD & POOR's, 2016), which is represented by ordinal rankings from obligors with the highest capacity to meet its financial obligations (rated as *AAA*) to those in default (rated as *D*). Throughout the estimations, as commonly employed for example in Avramov *et al.* (2007, 2013), ordinal numbers were applied to group stocks according to the level of credit quality of the firm. Thus, there were 22 levels of credit rating in the database, ranging from "*AAA*" to "*D*"; and the highest credit rating "*AAA*" is assigned the score 1 and to the lowest credit rating "*D*" the score 22[8]. The previous month credit ratings were assigned to the daily observations in the subsequent month, thus allowing an analysis of aggregate idiosyncratic variance by ratings using both daily and monthly frequency.

The analysis for the United States was divided into investment grade (firms with numerical score until 10, i.e BBB- or better credit ratings), non-investment grade (firms with numerical score equal or greater than 11, i.e, BB+ or worse credit ratings), and all rated firms. For the Brazilian case, the same procedure was employed, except that the investment grade division could not be done due to the non-availability of rated firms in this category for the beginning of the sample. Therefore, only two subsamples were considered.

Lastly, concerning the Brazilian data, CSV is broken according to the stock's market capitalization. Unlike the procedure employed by Angelidis and Tessaromatis (2008), which uses 80% and 20% of the largest and smallest stocks, respectively, here was adopted the median as a breakpoint as per the main procedure for construction of risk factors, specifically the size factor, as well as because in contrast with other developed countries (such as the UK and Australia) which have investigated and proposed different percentiles to obtain the risk factors; in Brazilian stock market these analyses are still incipient.

Concerning the analysis at the firm-level, monthly return observation greater than 300% were excluded from the sample as in Fu (2009) and Fink, Fink, and He (2012). A monthly frequency was adopted as the standard procedure in this literature and, as done by Fu (2009), and the smallest and largest 0.5% of all the control variables except of portfolio beta (Beta variable) were winsorized, i.e. replaced by their next closest value.

3.3.2 Construction of risk factors

It was necessary to estimate all the factors used in the multifactor models. The portfolios which were used to compute the size, value, momentum, operational profitability and investment factors were constructed on June 30th every year. When selecting the sample of stocks to be included in the final sample in a given year and conditioning in the model, the following data were required:

[8] Following Avramov *et al.* (2007, 2009), we apply this ordinal scale: *AAA* = 1, *AA*+ = 2, *AA* = 3, *AA*- = 4, *A*+ = 5, *A* = 6, *A*- = 7, *BBB*+ = 8, *BBB* = 9, *BBB*- = 10, *BB*+ = 11, *BB* = 12, *BB*- = 13, *B*+ = 14, *B* = 15, *B*- = 16, *CCC*+ = 17, *CCC* = 18, *CCC*- = 19, *CC* = 20, *C* = 21 and *D* = 22.

i. Positive stock market capitalization on the preceding June 30th for computing the size factor;
ii. Positive stock market capitalization and book value of equity on the preceding December 31st for computing the value factor;
iii. Non-missing prices on the preceding May 31th and July 31st for constructing the momentum factor;
iv. Non-missing total asset on the last two preceding December 31st for computing the investment factor;
v. Non-missing information for the revenues minus cost of goods sold, interest expense and selling, general and administrative expenses, and book value of equity on the preceding December 31st for computing the operational profitability factor.

Additionally, following the standard procedure, financial institutions and firms with negative book value were excluded from the construction of the risk factors. It is important to note that this research differs from Mussa, Rogers, and Securato (2009) and Mussa, Famá, and Santos (2012) in that the Brazilian factors are formed. The latter uses the market capitalization and book-to-market ratio on the company level to construct the factors, meaning that the common and preferred stocks are aggregated for each company. However, due to the small number of shares listed on the Brazilian stock exchange, the information used here was based on the share level, meaning that the common and preferred stocks are treated as two different assets implying that, for instance, two stocks of the same company may be in different percentiles and then be used to compose different risk factors.

The first factor, the market risk premium, MKT, was computed as the difference between the daily return on the Ibovespa index and daily return on the proxy for the risk-free asset (CDI). The construction of the SMB, HML, MOM, RMW, and CMA factors was similar to that suggested by Fama and French (1993, 2015) and in Kenneth French's website. The value-weighted portfolios were formed at the end of June each year and re-estimated annually, i.e. every 12 months. The weights were distributed among stocks according to their market capitalization and were updated monthly.

The steps to estimate the three-factor Fama and French (1993) model were:

i. At the end of June each year, the sample of eligible stocks was divided into two groups based on the median market capitalization;
ii. Within each of the two *size* groups, the stocks were divided into three groups according to the book-to-market ratio (as of the preceding December), the groups were split based on the 30^{th} and 70^{th} percentiles.

This method gives us 6 value-weighted portfolios for computing the empirical factors: 3 portfolios of small stocks and 3 portfolios of large stocks;

The size factor, SMB, is defined as the difference between the average monthly return on the three portfolios of small stocks and the average monthly return on the three portfolios of large stocks. The value factor, HML, is computed as the difference between the average monthly return on the two portfolios of high B/M ratio stocks and the average monthly return on the two portfolios with low B/M ratio stocks.

The steps to estimate the Carhart (1997) model were:

i. Repeating step i to obtain the three-factor model;
iii. Within each of the two *size* groups, the stocks were ordered based on the past 11-months accumulated returns (-12 months until -1 month) and divided into three groups based on the accumulated return using the 30^{th} and 70^{th} percentiles.

This method gives us 6 value-weighted portfolios: 2 portfolios of stocks with high prior return, 2 portfolios of stocks with medium prior return and, finally, 2 portfolios of stocks with low prior return.

The momentum factor, MOM, is defined as the difference between the average monthly return on the two portfolios of stocks with high prior return and the average monthly return on the two portfolios of stocks with low prior return.

The steps to estimate the five-factor Fama and French (2015) model were:

i. Repeating step i to obtain the three-factor model;
ii. Within each of the two *size* groups, the stocks were divided into three groups according to the book-to-market ratio, to the operational profitability (both as of the preceding December), and to the asset's growth (as of the last two preceding

Decembers). The groups were split based on the 30^{th} and 70^{th} percentiles. These procedures were necessary to obtain only the SMB factor;

iv. Within each of the two *size* groups (step i), the stocks were divided into three groups according to the book-to-market ratio (as of the preceding December), the groups were split based on the 30^{th} and 70^{th} percentiles;

v. Within each of the two *size* groups (step i), the stocks were divided into three groups according to the firm's asset growth (as of the two preceding Decembers). The groups were split based on the 30^{th} and 70^{th} percentiles;

vi. Within each of the two *size* groups (step i), the stocks were divided into three groups according to the operating profitability (as of the preceding December), the groups were split based on the 30^{th} and 70^{th} percentiles.

This method gives us the following value-weighted portfolios for computing the empirical factors: 9 portfolios of small stocks and 9 portfolios of large stocks; 2 portfolios of high B/M ratio stocks, 2 portfolios of low B/M ratio stocks; 2 portfolios of conservative investment stocks and 2 portfolios of aggressive investment stocks; and 2 portfolios of robust operating profitability stocks and 2 portfolios of weak operating profitability stocks.

The size factor, SMB, is defined as the difference between the average monthly return on the nine portfolios of small stocks and the average monthly return on the nine portfolios of large stocks. The value factor, HML, is computed as the difference between the average monthly return on the two portfolios of high B/M ratio stocks and the average monthly return on the two portfolios with low B/M ratio stocks. The conservative minus aggressive factor, CMA, is calculated as the difference between the average monthly return on the two portfolios of conservative stocks and the average monthly return on the two portfolios of aggressive stocks. The robust minus weak factor, RMW, is calculated as the difference between the average monthly return on the two portfolios of robust stocks and the average monthly return on the two portfolios of weak stocks.

3.3.3 Construction of Realized Idiosyncratic Volatility (IVOL) Portfolios for the Analysis of Aggregate Idiosyncratic Variance

It was also tested if return of quintiles portfolios sorted on realized idiosyncratic volatility (IVOL) as defined in Ang *et al.* (2006) can be predicted by cross-sectional variance. Guo and Savickas (2006) verify if Goyal and Santa-Clara's (2003) measure has predictive power for IVOL, while the interest here relies upon the chosen proxy for idiosyncratic variance. The procedure to create these portfolios was defined as: each month, all stocks with at least 15 daily observations were used and each stock returns within month were regressed on the Fama and French (1993) factors in order to estimate the residuals. Finally, stocks were sorted based on their realized IVOL in month *t*-1 and returns on these portfolios were calculated for the subsequent month. Stocks with the highest (lowest) IVOL were allocated into portfolio "High" ("Low"). Also, equal- and value-weighted portfolios were estimated, in addition to the value-weighted scheme employed by Guo and Savickas (2006).

3.3.4 Portfolio Formation for the Analysis of Expected Idiosyncratic Volatility at the Firm-Level

The abnormal returns relative to the model for which the forecasted idiosyncratic volatility variable was constructed (FF-3, FFC, FF-5) were estimated from a portfolio formation procedure based on the forecasted idiosyncratic volatilities (*Eivol*) using the EGARCH approach. It is interesting to note that the portfolio analysis allows the relation between variables to be investigated without limiting their functional interactions, but also restricts the number of variables employed at the same time, as it is not possible to control for them all together (FINK, FINK, HE, 2012).

In this sense, value-weighted portfolios with a monthly re-balancing scheme were created based only on a sorting on the *Eivol*. A double-sorting procedure was also carried out following Ang *et al.*'s (2006) approach, in which stocks were first sorted on a control variable (the present study used market capitalization, book-to-market, momentum and return reversal) and then according to the *Eivol*. In the univariate case, stocks were grouped into five groups, while in the double sorting procedure stocks were first allocated into two portfolios based on the median of the control variable (high and low values in the control variable), then within each portfolio stocks were allocated into three groups based on the *Eivol*, which resulted in 6 portfolios. The alphas within each *Eivol* category were averaged over each of the characteristic portfolios.

These analysis follows those implemented in Fink, Fink, and He (2012) and extended to include the Fama and French (2015) five-factor model as well, giving more flexibility to the assumed distribution during the forecasting of the idiosyncratic volatility.

3.3.5 Fama – MacBeth regressions for the Analysis of Expected Idiosyncratic Volatility at the Firm-Level

To conduct the analysis, Fama and MacBeth (1973, hereafter FM) cross-sectional regressions were run to test the sign and magnitude for forecasted idiosyncratic volatility when using a set of control variables. The first step was to replicate Fu's (2009) specification, which includes the portfolio beta, market capitalization, book-to-market, momentum, turnover, coefficient of variation of turnover, and also including the short-term return reversal (REV), and *Eivol* based on three alternative asset pricing models in the mean equation with an EGARCH (1,1) model. These control variables are explained in the next section. The short-term return reversal (REV) was included to investigate the role of the lagged return in the expected returns and idiosyncratic volatility relation, as suggested by Huang *et al.* (2010) and followed by Fink, Fink, and He (2012). The cross-sectional regressions were run to test the sign and magnitude for *Eivol* were as follows:

$$R_{it} = \gamma_{0t} + \sum_{k=1}^{K} \gamma_{kt} X_{kit} + \varepsilon_{it} \qquad i = 1,2,\dots,N_t, \quad t = 1,2,\dots \qquad \text{Equation 5}$$

Where $\mathrm{R_{it}}$ represents the realized return on stock i in month t; X_{kit} refers to the set of control variables employed in this study; ε_{it} is the error term; N_t specifies the number of stocks in each month t; and T is the total number of months in the sample (number of cross-sectional regressions carried out here), in this case, 143 (August 2004 to June 2016). The coefficients are then averaged over time and the t-statistic is computed to evaluate their significance. The formulas used for this purpose are presented below:

$$\hat{\gamma}_k = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_{kt} \qquad \text{Equation 6}$$

$$t_k = \frac{\widehat{\gamma_k}}{std_{\hat{\gamma}_k}/\sqrt{T}} \qquad \text{Equation 7}$$

Where $std_{\hat{\gamma}_k}$ is the standard deviation of $\hat{\gamma}_k$. Aditionally, the traditional FM regression assumes an equal-weighted scheme; as a result, small and big firms enter as the same level, which can interfere in the estimations. In fact, Ang *et al.* (2009) affirm that this may add noise to the estimations by assuming the equal weight on small firms. In this sense, trying to verify the consistency of the results, Generalized Least Squares (GLS) regressions were also run using the market capitalization as a weighting scheme along the diagonal. It is important to note, as mentioned by Ang *et al.* (2009), that value-weighted FM regressions are similar to value-weighted portfolios and the equally-weighted to the equal-weight portfolios.

3.3.6 Model Validation

All regressions estimated were examined to verify any potential issues that could bias a correct interpretation of the results. Thus, the Ljung-Box (1978) test for autocorrelation and Breusch-Pagan (1979) test for heteroscedasticity were run in order to check for the presence of these characteristics. The functional form of these tests can be found in Greene (2002). The t-statistics reported referred to those Newey-West (1987) adjusted. The United States stock market, using aggregate idiosyncratic variance, was followed the same number of lags specified by Garcia, Mantilla-García, and Martellini (2014) while the Brazilian case at aggregate or individual level the suggestion in Davidson and Mackinnon (1993) was employed with a rule as $T^{1/4}$, where T is the number of observations. All estimations were done using the R software.

3.4 Variables

This subsection details how the variables used in both the aggregate idiosyncratic risk and at the firm-level idiosyncratic volatility analysis were estimated in order to accomplish the objectives previously defined.

3.4.1 Aggregate Idiosyncratic Variance (Cross-sectional Variance)

The CSV formulation outlined here follows closely the presentation shown by Garcia, Mantilla-García, and Martellini (2014). First, let's assume that N_t is the number of stocks in period t in a portfolio and i = 1, 2,..., N_t. The return of each stock i in time t is represented by r_{it}, then the return on the portfolio $r_t^{(wt)}$, as defined by Markowitz (1952), is given by:

$$r_t^{(w_t)} = \sum_{i=1}^{N_t} w_{it} r_{it} \qquad \text{Equation 8}$$

Where: w_t is the weight vector of each stock in the portfolio.

The restrictions made by Garcia, Mantilla-García, and Martellini (2014) of using portfolios composed with at least two stocks and with positive weights in each time t are maintained. Thus, the cross-sectional variance measure, denoted by $CSV_t^{(w_t)}$, proposed by the authors using a given weighting scheme (w_t) is computed by:

$$CSV_t^{(w_t)} = \sum_{i=1}^{N_t} w_{it}(r_{it} - r_t^{(w_t)})^2 \qquad \text{Equation 9}$$

As usual in the context of portfolio selection, one of the two weighting schemes implemented by Garcia, Mantilla-García, and Martellini (2014) is the equal-weighted CSV, hereafter named CSV_t^{EW} where $w_{it=}1/N_t$ $\forall$ i and t. Thus, assuming that r_t^{EW} represents the return on the equal-weighted portfolio, it is shown that:

$$CSV_t^{EW} = \frac{1}{N_t} \sum_{i=1}^{N_t}(r_{it} - r_t^{EW})^2 \qquad \text{Equation 10}$$

In the same manner, the second weighting scheme used by the authors and based on the market capitalization is the value-weighted CSV, hereafter named CSV_t^{CW}. The total market capitalization (C_t) is obtained by summing up the market capitalization of each stock i (c_{it}) in the beginning of the month of the respective day t, and so $C_t = \sum_{i=1}^{N_t} c_{it}$; the CSV_t^{CW} is calculated by:

$$CSV_t^{CW} = \sum_{i=1}^{N_t} w_{it}^{CW}(r_{it} - r_t^{CW})^2 \qquad \text{Equation 11}$$

Where: $w_{it}^{CW} = c_{it}/C_t$ is the market capitalization weighting scheme and r_t^{CW} is the return on the market capitalization-weighted portfolio.

Garcia, Mantilla-García, and Martellini (2014) document a formal investigation of the behavior of their proposed CSV measure under the asymptotic theory and the assumptions of homogeneous betas ($\beta_{it} = \beta_t = 1\ \forall\, i$) and residual variance ($\mathrm{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2(t)\ \ \forall\ \ i$). The authors describe that the $CSV_t^{(w_t)}$ is biased by $(1 - \sum_{i=1}^{N_t} w_{it}^2)$, but in the case of CSV_t^{EW} it is the best estimator considering a strictly positive weighting scheme, in which when the number of firms increases infinitely its bias and variance is shown to be, respectively:

$$\mathrm{E}[CSV_t^{EW}] \xrightarrow[N_t \to \infty]{} \sigma_\varepsilon^2(t)$$

$$var(CSV_t^{EW}) \xrightarrow[N_t \to \infty]{} 0$$

But the same does not happen to the market capitalization (CW) weighting scheme. When relaxing the assumption of homogeneous beta, an extra positive bias is introduced in the $CSV_t^{(w_t)}$ and it remains even in the case of an infinite number of firms. This aditional bias is given by $\mathrm{E}\left[F_t^2 CSV_t^\beta\right]$, where F_t^2 represents the average return of the market portfolio squared, and CSV_t^β is the cross-sectional variance of stock betas[9].

The total bias is obtained as the intercept of the regression of the $CSV_t^{(w_t)}$ with both EW and CW weighting schemes on the average idiosyncratic variance obtained using the CAPM and FF-3 models. Their empirical results show that for both EW and CW, using the CAPM and FF-3 as model, the bias is in the order of 10^{-5}, then empirically it does not affect estimations.[10]

[9] The cross-sectional dispersion in betas in Garcia, Mantilla-García, and Martellini (2014) is given by: $CSV_t^\beta = \sum_{i=1}^{N_t} w_{it}\,(\beta_{it} - \sum_{j=1}^{N_t} w_{jt}\beta_{jt})^2$, where the beta is measured by ratio of the covariance between the asset's return and the market return and the variance of the market return.

[10] Table 2 presented in Garcia, Mantilla-García, and Martellini (2014, p. 1143), which verifies the size of the bias in their analysis, was replicated using Brazilian stocks and the bias found was close to those reported by the authors (around 10^{-5}), then it was also believed that though it is significant it still has no practical influence on the results. The maximum and minimum number of stocks in the Brazilian case is 281 and 104, respectively.

3.4.2 Alternative Measures of Aggregate Idiosyncratic Variance

The CSV measure is interpreted as a proxy to idiosyncratic risk as it was shown to be highly correlated with some other widely competitive measures by Garcia, Mantilla-García, and Martellini (2014). Thus, this subsection describes the four alternative measures which were calculated for the Brazilian case to verify if the same pattern exists.

- Multifactor models – Ang *et al.* (2006, 2009). This first measure is based on the residuals of multifactor models. In this case, this works apply the following models to obtain the estimated residuals: CAPM, three-factor Fama and French (1993), Carhart (1997), and, more recently, the five-factor Fama and French (2015) model, which adds the investment and operational profitability risks premia.

  In this case, r_{it} is the return of stock *i* at time *t*, r_f represents the risk-free asset, $MKT_t, SMB_t, HML_t, Mom_t, RMW_t$, and CMA_t are the excess market return, the small minus big factor is based on the market capitalization of stocks, the high minus low factor is based on the book-to-market ratio, the robust minus weak factor is based on the firm's operating profitability and the conservative investment minus aggressive factor is based on the firm's assets growth, respectively. Then, the idiosyncratic variance of stock *i* in period t relative to these models, $\sigma^2\left(\varepsilon_{it}^{model}\right)$ is based on the variance of the residuals of the following models:

$$r_{it} - r_{ft} = \alpha_{it} + \beta_{it} MKT_t + \varepsilon_{it}^{CAPM} \quad \text{Equation 12}$$

$$r_{it} - r_{ft} = \alpha_{it} + \beta_{it} MKT_t + s_{it} SMB_t + h_{it} HML_t + \varepsilon_{it}^{FF3} \quad \text{Equation 13}$$

$$r_{it} - r_{ft} = \alpha_{it} + \beta_{it} MKT_t + s_{it} SMB_t + h_{it} HML_t + m_{it} MOM_t + \varepsilon_{it}^{FFC} \quad \text{Equation 14}$$

$$r_{it} - r_{ft} = \alpha_{it} + \beta_{it} MKT_t + s_{it} SMB_t + h_{it} HML_t + rw_{it} RMW_t + c_{it} CMA_t + \varepsilon_{it}^{FF5}$$

Equation 15

The estimated average idiosyncratic variance, as in Bekaert, Hodrick, and Zhang (2012) with a market capitalization weighting is then:

$$model_t^{CW} = \sum_i^{N_t} w_{it} \sigma^2\left(\varepsilon_{it}^{model}\right) \quad \text{Equation 16}$$

It is noteworthy that Equation 16 was also estimated using an equal- weighted scheme as in Garcia, Mantilla-García, and Martellini (2014).

- Goyal and Santa- Clara (2003) – r_{id} is the return of the stock *i* in day d and D_t is the number of days in month *t*; the Goyal and Santa-Clara (2003) average stock variance, which is mostly idiosyncratic, is given by:

$$GS_t^{EW} = \frac{1}{N_t}\sum_{i=1}^{N_t}\left[\sum_{d=1}^{D_t} r_{id}^2 + 2\sum_{d=2}^{D_t} r_{id}r_{id-1}\right] \qquad \text{Equation 17}$$

Then, as can be seen, this is another measure which uses daily returns within a month to produce monthly estimates of the average stock variance, but the second term of the equation allows for adjustments for the autocorrelation that can exist in the returns according to French, Schwert, and Stambaugh (1987). In some cases, the second part made the estimation negative and was then ignored for these cases (GOYAL; SANTA-CLARA, 2003). The CW scheme was obtained as well.

Additionally, the difference between GS_t^{EW} or GS_t^{CW}, which uses a market capitalization weighting scheme, and the corresponding market variance was also calculated as in Jiang and Lee (2006) and Angelidis and Tessaromatis (2008), named here as $GS - M^{EW}$, when using an equal-weighted distribution, and $GS - M^{CW}$, with a market capitalization scheme. In other words, this additional measure employed here subtracts the market variance from the total variance in order to estimate idiosyncratic variance. When this measure was estimated, market variance was given by summing up the squared market returns adjusted for autocorrelation within the month (see Goyal and Santa-Clara, 2003, and Jiang and Lee, 2006).

- Bali, Cakici and Levy (2008) – this measure is based on the gain from portfolio diversification and, as stated by the authors, does not require beta estimations nor correlations, being characterized as a model-free measure.

$$\sigma_{\varepsilon,t}^2 = \left(\sum_{i=1}^{n} w_{i,t}\sigma_{i,t}\right)^2 - var\left(R_{m,t}\right) \qquad \text{Equation 18}$$

Where: $\left(\sum_{i=1}^{n} w_{i,t}\sigma_{i,t}\right)^2$ represents the variance of the nondiversified portfolio whereas $var\left(R_{m,t}\right)$ corresponds to the variance of the diversified portfolio (market variance).

3.4.3 Equal- and Value-weighted Market Returns

The United States's daily and monthly market returns corresponded to the equal- and value-weighted CRSP indexes. For the Brazilian case, daily market returns were composed from all stocks in the sample and the market capitalization used was the one at the beginning of the month. Monthly market returns were obtained by simply accumulating the returns over the month.

3.4.4 Market Variance

In accordance with Garcia, Mantilla-García, and Martellini (2014), an EGARCH was applied to daily equal- and value-weighted market return series to obtain conditional variance; for the United States stock market, an EGARCH (1,1) was adopted for both equal- and value-weighted market returns whereas, for the Brazilian case, EGARCH (3,3) was used in both cases as selected by the best model with minimum Akaike Information Criteria (AIC). Monthly variance was calculated as the monthly realized market variance.

3.4.5 Cross-sectional skewness (CSS) and kurtosis (CSK)

Garcia, Mantilla-García, and Martellini (2014) analyzed if a robust measure of skewness based on quintiles defined in Bowley (1920), generalized by Hinkley (1975) and used in Kim and White (2004), namely the cross-sectional skewness (CSS), is useful in predicting market returns: with exception of the monthly and value-weighted scheme, all the remaining analysis showed that CSS plays a role in forecasting aggregate returns. The corresponding analysis were employed here and extended to include the cross-sectional kurtosis (CSK).

$$CSS = \frac{F^{-1}(1-\alpha_1) + F^{-1}(\alpha_1) - 2Q_2}{F^{-1}(1-\alpha_1) - F^{-1}(\alpha_1)} \qquad \text{Equation 19}$$

For any α_1 between 0 and 0.5 and $Q_2 = F^{-1}(0.5)$ and F^{-1} represents the inverse of the cumualtive distribution function (c.d.f). Kim and White (2004) add that when $\alpha_1 = 0.25$, Bowley's (1920) measure become a special case of Hinkley's (1975) generalization and that this measure satisfies Groeneveld and Meeden (1984) propositions of what corresponds to an adequate measure of skewness. In relation to cross-sectional kurtosis one robust measure studied in Kim and White (2004) and originally presented in Crow and Siddiqui (1967) was again followed, which, in turn, is also based on quintiles. Its definition is as follows:

$$CSK = \frac{F^{-1}(1-\alpha_1) + F^{-1}(\alpha_1)}{F^{-1}(1-\beta_1) - F^{-1}(\beta_1)} \qquad \text{for } \alpha_1, \beta_1 \in (0,1) \qquad \text{Equation 20}$$

In this case, the authors used the values 0.025 and 0.25 for α_1 and β_1, respectively.

3.4.6 Expected Idiosyncratic Volatility at the Firm-Level

The EGARCH (*p*,*q*) model is used to obtain forecasted idiosyncratic volatility to use at the firm-level analysis, where *p* represents the number of lags related to the past conditional variance and *q* the number of lags of return shocks. It is part of a class or family of models that are continuously expanding in the econometric literature originally derived by Engle (1982) with the introduction of the Autoregressive Conditional Heteroscedastic (ARCH), and later extended by Bollerslev (1986), called the Generalized Autoregressive Conditional Heteroskedastic (GARCH). Both models are nonlinear in variance and take into account some of the characteristics of financial data, such as clustering, but whereas the former evaluates conditional volatility as dependent only of its past innovations, the latter also relates it to its past variance.

The GARCH model is able to capture volatility clusters. The usually assumed distribution is the normal distribution, which does not adequately represent the data when there is presence of skewness and kurtosis. Brooks *et al.* (2005) note that variance is a risk measure well used under the normality assumption, however if the returns have heavy tails (leptokurtosis), this will lead to an underestimation of the true risk of the portfolio. In addition, a critical aspect pointed out by Gibbons, Ross, and Shanken (1989) is that the increase in the number of assets in the portfolio does not diversify the asymmetry and kurtosis.

In relation to kurtosis, Mandelbrot (1963) and Fama (1965) were the first to observe leptokurtosis in the data, which later became a characteristic exhibited by the financial series in their great majority (CONT, 2001). According to Bai, Russell, and Tiao (2003), leptokurtosis can be induced due to both volatility clustering and conditional non-normality shown in the data. Brooks *et al.* (2005) synthesizes that leptokurtosis implies that extreme movements (up or down) will occur at a higher frequency than expected by the gaussian distribution.

Skewness and kurtosis represent the third and fourth moment of the distribution, respectively. In this sense, skewness can be positive or negative: in the first case, there is a greater probability of positive returns than negative returns of the same magnitude after the mean is subtracted; in the second case, the opposite occurs, i.e. a greater probability of negative returns than positive ones of the same magnitude after subtracting from the mean (HARVEY, SIDDIQUE, 1999). In addition to this, kurtosis represents an increase in the probability of occurrence of extreme events (FANG, LAI, 1997).

The literature sheds some light on why the asset's return can be asymmetric. Damodaran (1985) suggests it is related to the dissemination of good and bad news about the company in the market; Chen, Hong, and Stein (2001) relates it to the heterogeneity among investors, and Bae, Lim, and Wei (2006) mention the differences in the quality of corporate governance. Here, a more flexible distribution for the error term was employed.

It is noteworthy that GARCH is also a symmetric model, meaning it does not distinguish between positive and negative impacts of the same magnitude of returns. However, empirical tests often show that this does not occur in practice. The leverage effect initially studied in Black (1976) and Christie (1982) shows that negative shocks have greater influence on volatility than positive shocks of the same magnitude. This represents a different response (asymmetry) in the conditional variance depending on the sign of the returns. As a consequence, Nelson (1991) developed the Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH), which considers the differences of shocks of the same magnitude over variance.

The EGARCH (1,1) is defined as follows:

$$r_{it} - r_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + \varepsilon_{it}^{FF3} \quad \text{Equation 21}$$

$$r_{it} - r_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + m_i MOM_t + \varepsilon_{it}^{FFC} \quad \text{Equation 22}$$

$$r_{it} - r_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \varepsilon_{it}^{FF5}$$

Equation 23

$$\varepsilon_t \sim SGED(\kappa, \nu)$$

$$log\sigma_{it}^2 = \varpi_i + b_i z_{t-1} + \gamma_i(|z_{t-1}| - \mathrm{E}|z_{t-1}|) + c_i log\sigma_{t-1}^2$$ Equation 24

Where: $t = 1, \ldots, T$ and $i = 1, \ldots, N$, ϖ_i is a constant of the model, b_i represents a sign effect; γ_i refers to a size effect or a magnitude change parameter; c_i represents the effect of past variance; z_t corresponds to the standardized innovation; and $\mathrm{E}|\varepsilon_t|$ refers to the expectation of the absolute value of the error term based on a Skew-GED distribution. κ and ν are the shape thickness parameter and an asymmetry parameter, respectively.[11] Again, FF -3 represents the Fama and French (1993) three-factor model, FFC corresponds to the Carhart (1997) model and FF-5 to the Fama and French (2015) five-factor model.

Pagan and Schwert (1990) and Engle and Mustafa (1992) affirm that the EGARCH model could fit most of the asymmetry phenomenon in the U.S. and Japan stock markets, and also in the stock options market using American data. This model was applied by Fu (2009), Brockman, Schutte, and Yu (2009), Fink and Fink, and He (2012), being used here as well. These authors have used a combination of p,q GARCH/ARCH lags varying from 1 to 3 based on the minimum Akaike Information Criteria (AIC).

Here only the EGARCH (1,1) was adopted using all the information available to traders to forecast the idiosyncratic volatility, a Skew-GED distribution to accommodate non-normality and a setting with 2.000 iterations as done in Fink, Fink, and He (2012) and trying to reduce the hypothesis of restricting this value too much as Fu (2010) argues about Guo, Kassa, and Ferguson's (2014) paper. An expanding window requiring a minimum of 60 data points was employed and the model was re-estimated with every new information added to the window.

[11] For more details in the Skew-Ged functional form, see Fernandez and Steel (1998).

3.4.7 Illiquidity ratio

A liquidity measure was included to check if idiosyncratic variance can be seen as a proxy for liquidity risk. Angelidis and Tessaromatis (2008) describe that Amihud's (2002) liquidity was found to be the best proxy to obtain liquidity as in Hasbrouck (2005) and Acharya and Pedersen (2005), as it exhibits a high degree of correlation with other measures constructed from microstructure data. Then, the authors include this measure in their analysis of idiosyncratic variance. Bali et al. (2005) also use the liquidity measure when testing the validity of Goyal and Santa – Clara (2003) results regarding the importance of idiosyncratic variance to predict market returns. In this sense, the illiquidity aspect was also investigated in the current study, where I represents the illiquidity, $r_{i,d}$ is the daily return on stock *i* on day *t*, $Vol_{i,t}$ is the trading dollar volume on stock *i* on day *t*, and *n* is the number of observations, then:

$$I_{i,t} = \sum_{t=1}^{n}\left\{\left|r_{i,d}\right|/Vol_{i,t}\right\}/n \qquad \text{Equation 25}$$

This measure is interpreted in the sense that the higher $I_{i,t}$ is, then the higher the stock's illiquidity is, where returns are changed significantly, although not with a significant volume. To obtain an aggregate illiquidity estimate for the whole market, $I_{i,t}$ is simply averaged across the stocks. After that, the log of aggregate illiquidity was regressed on its past value (*t*-1) to decompose it into expected $(Illq_t^{\mathrm{E}})$, from the predicted values, and unexpected components $\left(Illq_t^{\mathrm{U}}\right)$, from the residuals, measures of illiquidity. Amihud (2002) posits that expected returns have a positive (negative) association to expected illiquidity (unexpected illiquidity).

3.4.8 Portfolio Beta

The procedure outlined in Fama and French (1992) and repeated by Fu (2009) and Brockman, Schutte, and Yu (2009) was replicated to obtain the portfolio's Beta, instead of directly using the asset's beta by itself to represent systematic risk. Every month, the 60 previous months' returns were used to estimate the stock beta and then sorted into two portfolios by size and within each portfolio stocks were sorted on three portfolios by their beta (2x3 sorting procedure). Equal-weighted portfolio returns were constructed each month according to the double sorted portfolios.

After that, a time series regression for each of the portfolio returns on the current and lagged month market return were performed. The next step was to sum up the slopes of the two regression coefficients, known as the portfolio Beta (Beta), to adjust for non-synchronicity in trading. This value is assigned to every stock according to the portfolio it belongs to across the months. Fu (2009) uses a 10x10 sorting procedure while Brockman, Schutte, and Yu (2009) use a maximum of 6 portfolios to some countries due to the number of stocks. This explains why this 2x3 procedure was chosen and why portfolios with a very reduced number of stocks were avoided. The mean beta was 0.89 and the median was 0.84.

3.4.9 Market capitalization

Market value (MV) was calculated for each stock by multiplying its monthly price and the number of outstanding shares. During the Fama-MacBeth (1973) regressions estimations, its lagged value is used following Huang *et al.* (2010). The natural log is taken for MV, denoted as LNMV.

3.4.10 Book-to-market

The book-to-market (B/M) relation, which is the inverse of the market-to-book ratio, was obtained from the firm's book value as given in the Quantum® database divided by the stock market value. As in Fu (2009), the procedure described in Fama and French (1992) was applied: book value of December t-1 was divided by the market value in December t-1. This ratio was matched to the book-to-market relation from July of year t to June of year t+1. Every year, the procedure is repeated to ensure that information was assimilated by the market. The natural log is taken for B/M, denoted as LNBM.

3.4.11 Momentum

Following Fu (2009), Brockman, Schutte, and Yu (2009), and Fink, Fink, and He (2012), the variable momentum, named as RET(-2,-7), refers to past returns and was obtained from month t-7 to t-2, where t refers to the month of expected return. As Fu (2009), it was obtained the compound gross return and the one-month lag return was not used to mitigate interferences of thin trading or bid-ask spread.

3.4.12 Turnover

Following Fu (2009), Brockman, Schutte, and Yu (2009), and Fink, Fink, and He (2012), the turnover variables were defined as in Chordia, Subrahmanyam, and Anshuman (2001) to represent the stock’s liquidity during the analysis at the firm-level idiosyncratic volatility. As in Fu (2009), the average turnover (TURN) from the past 36 months was used: the turnover in a given month is the division of number of shares traded by the number of shares outstanding. The coefficient of variation of turnover (CVTURN) for the past 36 monthly was also obtained to control for liquidity. For both TURN and CVTURN the natural logs were taken denoted as LNTURN and LNCVTURN.

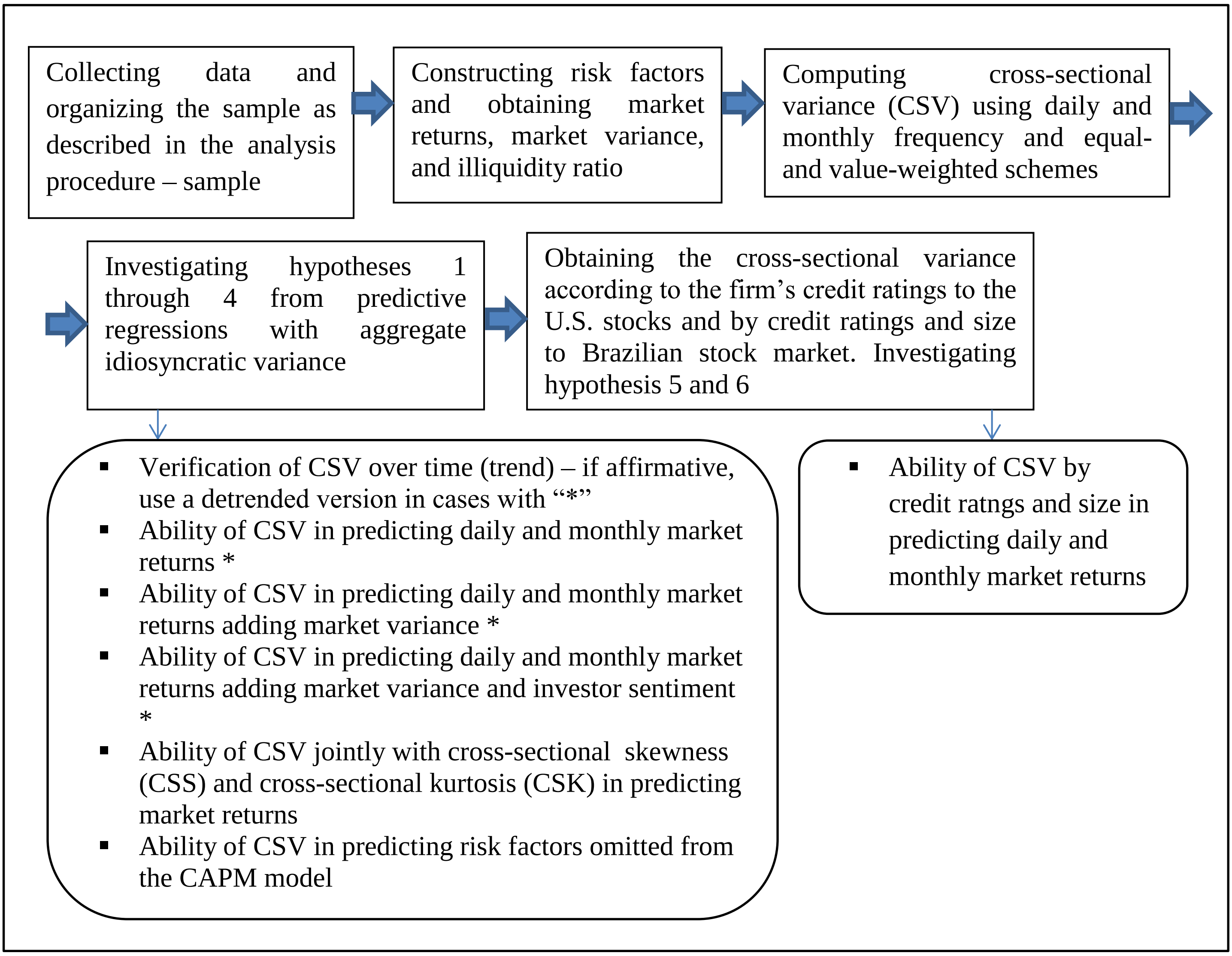
3.4.13 Short term reversal

Following Fu (2009), Huang *et al.* (2010), and Fink, Fink, and He (2012), the role of return reversal (REV) was also considered, which represents the lagged one-month return.

3.5 Research Scheme

In order to have a better view of the procedures employed in this thesis, Figure 1 shows the steps followed to obtain the main results concerning the relation between aggregate idiosyncratic variance and market returns.

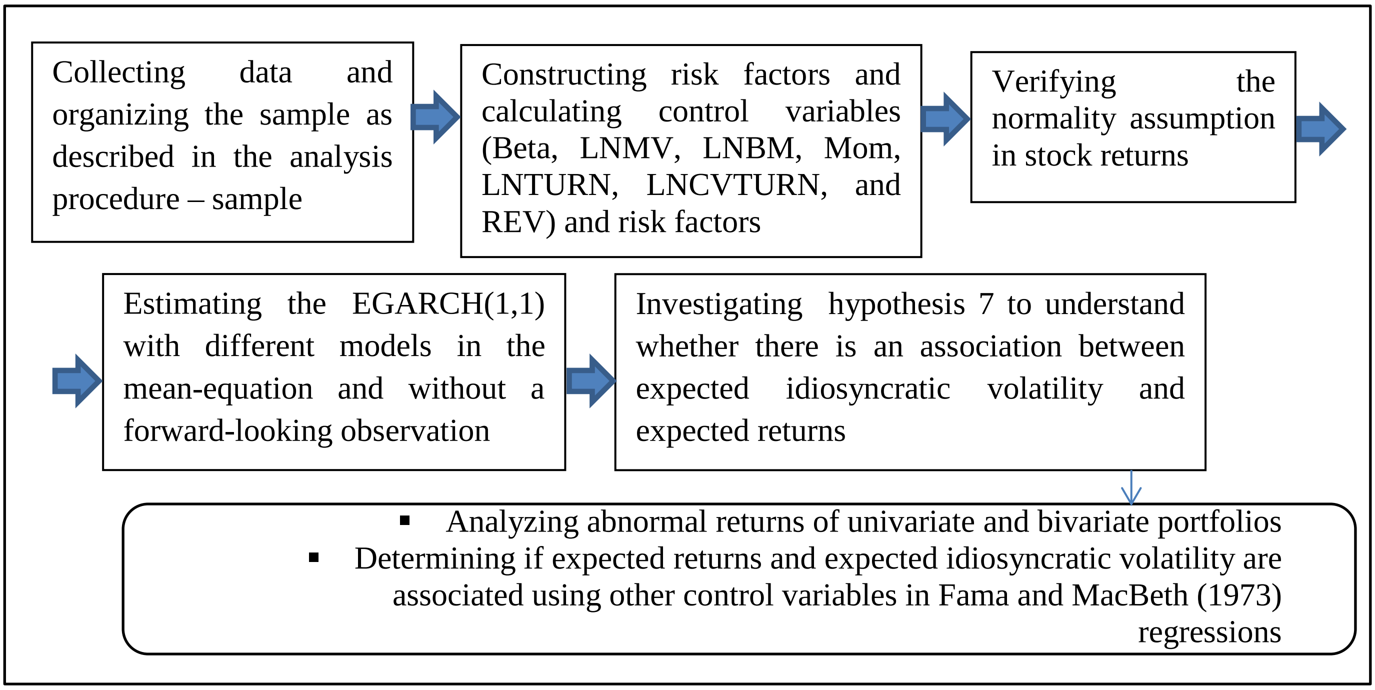
Figure 1 - Overview of the Steps followed with Aggregate Idiosyncratic Variance



Source: Author (2017)

In the same manner, Figure 2 describes how the procedures to obtain the analysis of individual idiosyncratic volatility and expected returns were followed in this thesis.

Figure 2 - Overview of the Steps Followed with Expected Idiosyncratic Volatility at the Firm-Level



Source: Author (2017)

4 EMPIRICAL RESULTS

This section presents and discusses the main results found in the current research. First, the general findings concerning the behavior of idiosyncratic variance are investigated and then its importance to predict market returns is analyzed controlling for a variety of model specifications. Moreover, its ability to predict risk factors omitted from the CAPM model and quintiles portfolios formed according to the stock's realized idiosyncratic volatility is investigated. Next, the construction of aggregate idiosyncratic variance according to the firm's credit ratings and size are considered to understand if the idiosyncratic risk of stocks with these characteristics helps predict market returns. Lastly, the individual effects of idiosyncratic volatility are analyzed by investigating the relationship between expected idiosyncratic risk and expected returns.

4.1 Idiosyncratic Variance at the Aggregate Level

4.1.1 Descriptive Statistics of Aggregate Idiosyncratic Variance

The general behavior of the cross-sectional variance using both equal- and value-weighted market capitalization in daily and monthly frequency is first shown in Table 2.

Table 2 - Daily and Monthly Measures of Idiosyncratic Variance – United States Sample

Annualized mean and standard deviation of the CSV measures proposed by Garcia, Mantilla-García, and Martellini (2014). Panel 1A shows the daily measures using the same sample period originally investigated by the authors from January 1964 to December 2006 and also extended to December 2014. Panel 2B shows the results for the monthly frequency. Panel 1B and 2B reports their corresponding correlation matrixes.

1964:01 - 2006-12			1964:01 - 2014:12		
Panel 1A: Daily					
	CSV^{CW}	CSV^{EW}		CSV^{CW}	CSV^{EW}
Mean	0,092	0,433	Mean	0,092	0,434
Standard deviaton	0,005	0,021	Standard deviaton	0,006	0,028

Continued

Panel 1B: Correlation Matrix of Daily Measures					
	CSV^{CW}	CSV^{EW}		CSV^{CW}	CSV^{EW}
CSV^{CW}	100	56,02	CSV^{CW}	100	51,49
CSV^{EW}		100	CSV^{EW}		100
Panel 2A: Monthly					
	CSV^{CW}	CSV^{EW}		CSV^{CW}	CSV^{EW}
Mean	0,092	0,433	Mean	0,092	0,433
Standard deviaton	0,020	0,084	Standard deviaton	0,021	0,095
Panel 2B: Correlation Matrix of Monthly Measures					
	CSV^{CW}	CSV^{EW}		CSV^{CW}	CSV^{EW}
CSV^{CW}	100	66,63	CSV^{CW}	100	69,27
CSV^{EW}		100	CSV^{EW}		100

Source: Author (2017)

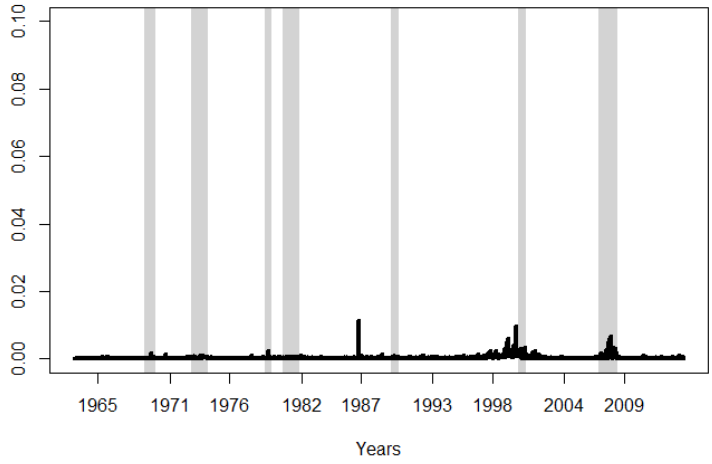
From both measures using daily and monthly frequency it can be observed that the equal-weighted scheme generates higher values than those showed to the value-weighted idiosyncratic variance. In the extended sample period, the annualized mean and standard deviation remain virtually unchanged, most probably by the number of stocks in the sample, but the correlation between equal and value-weighted daily CSV measures decreases and at the monthly frequency it has a small increment.

Considering the period until 2006, the results shown in Table 2 are very close to those reported in Garcia, Mantilla-García, and Martellini's (2014) paper. Figure 3 displays daily idiosyncratic variance over time and the shaded area represents periods of recession as defined by the NBER dates.

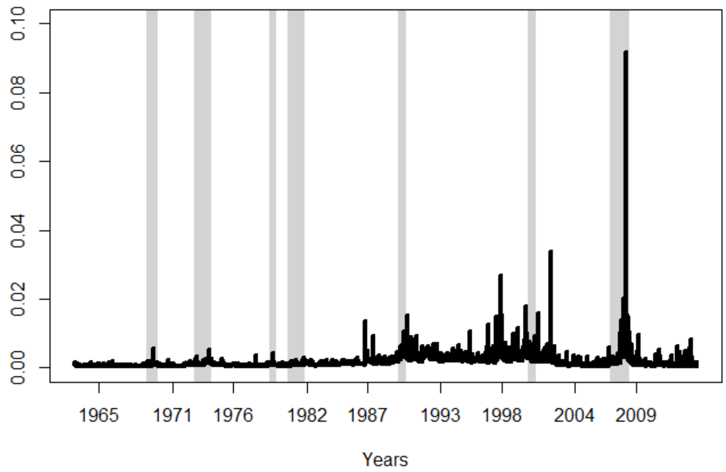
Figure 3 - Daily Idiosyncratic Variance – United States Sample

The figure displays the daily idiosyncratic variance for the United States stock market, where: a) represents the value-weighted measure; and b) the equal-weighted version. The sample period is from January 1964 to December 2014.

a)



b)



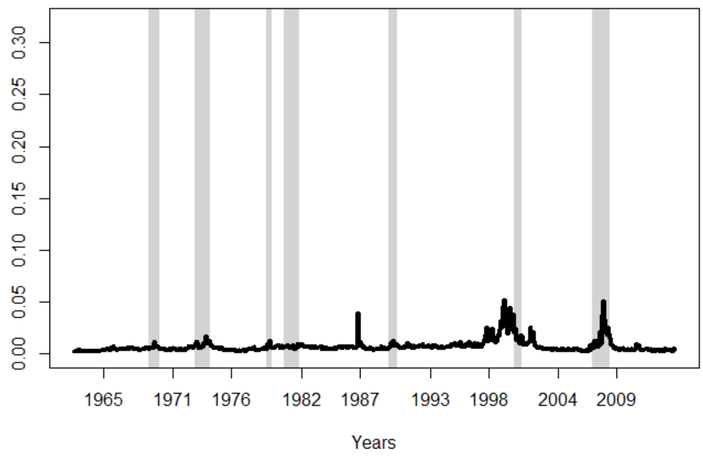
Source: Author (2017)

As shown in Table 2, CSV^{EW} has higher values than those of CSV^{CW}. In fact, in the period of 1987, the latter achieves its maximum while for the former it occurs only during the 2007/2008 subprime crisis, showing they have been influenced by extreme downturns in different ways. Figure 4 depicts the monthly CSV.

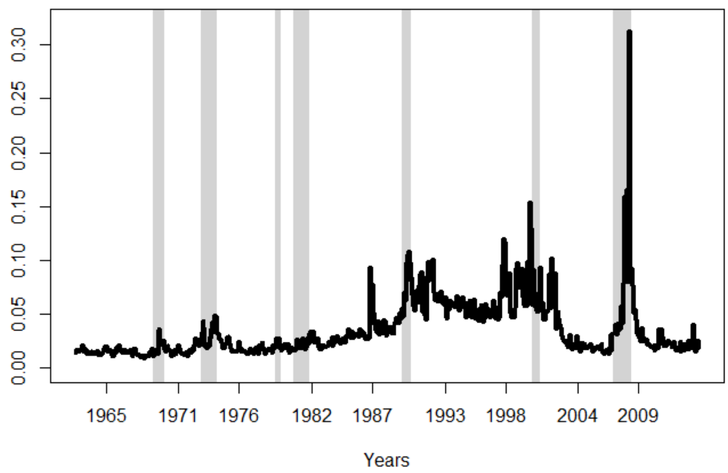
Figure 4 - Monthly Idiosyncratic Variance – United States Sample

The figure displays the monthly idiosyncratic variance for the United States stock market where a) represents the value-weighted measure and b) the equal-weighted version. The sample period is from January 1964 to December 2014.

a)



b)



Source: Author (2017)

With a monthly frequency, CSV^{CW} shows its maximum around April 2000 and October 2008. The first episode corresponds to the Internet companies (dotcom) bubble and the period where Garcia, Mantilla-García, and Martellini (2014) affirm to be their maximum number of firms. The second moment illustrates the impact of the subprime crisis on its monthly series.

In contrast, CSV^{EW} remains strongly affected by the 2007/2008 crisis reflecting mainly the behavior of the increased risk in small firms.

Using Brazilian stocks, CSV was constructed to understand the behavior of idiosyncratic variance from an emergent country's point of view. Table 3 exhibits its main behavior along the years for daily frequency.

Table 3 - Daily Measure of Idiosyncratic Variance – Brazilian Sample

The table presents the annualized mean and standard deviation of the daily CSV measures proposed by Garcia, Mantilla-García, and Martellini (2014) and their correlation. The sample period is from January 2000 to June 2016.

Panel A: Mean and Standard Deviation of Idiosyncratic Variance		
	CSV^{CW}	CSV^{EW}
Mean	0,109	0,443
Standard deviaton	0,009	0,042
Panel B: Correlation Matrix between of Idiosyncratic Variance		
	CSV^{CW}	CSV^{EW}
CSV^{CW}	100	37,15
CSV^{EW}		100

Source: Author (2017)

Both measures of CSV (equal- and value-weighted) show a similar pattern in terms of annualized mean and standard deviation to the United States case using daily data throughout the timespan analyzed, however the correlation between them is much lower (37,15%) than the correlation seen in the that stock market (>50%). Consequently, it translates into a higher importance for the weighting scheme used and may reflect the characteristics of an emergent country with a reduced number of stocks raising a difference between small and big firms by their market capitalization.

Table 4 summarizes CSV and other competitive measures (Fama and French, 1993; Carhart, 1997; Goyal and Santa-Clara, 2003; Bali, Cakici, and Levy, 2008; Fama and French, 2015), to understand how well it represents idiosyncratic variance in a monthly frequency. It is interesting to observe that CSV has a similar mean and standard deviation to those shown here in the United States case and also with other competitive measures shown in Garcia, Mantilla-García, and Martellini's (2014) results.

Panel B of Table 4 shows the correlation matrix for the different idiosyncratic measures adopted. It is very interesting to observe that it exhibits a strong association between measures using the same weighting scheme and this result is important as it allows us to understand the cross-sectional variance as a proxy to idiosyncratic variance taking into account the mostly used measures in the literature. However, the correlation between the equal-weighted versus value-weighted versions of the same measure are substantially lower highlighting their differences.

Figure 5 displays daily idiosyncratic variance over time and the shaded area represents periods of recession as defined by the Business Cycle Dating Committee (Getulio Vargas Foundation - Brazilian Institute of Economy). Equal-weighted idiosyncratic variance exhibits higher values than value-weighted while both show a rise around the end of 2003, more specifically in August 2003 and the value-weighted idiosyncratic variance in October of the same year, which could be explained by political reasons given by the expectations about decisions to be adopted during the new government. Both measures also have a sharply increase during the 2007/2008 crisis reflecting a higher idiosyncratic risk.

Figure 6 depicts the monthly idiosyncratic variance. Again, monthly value-weighted idiosyncratic variance has lower values than its equal-weighted counterpart as well as being less volatile. Aggregate idiosyncratic variance using the equal-weighting scheme has much more volatility suggesting that small firms in the Brazilian case differ from the value-weighted version. The highest peak occurred in 2008 reflecting the subprime crisis which took place in the United States market and spreaded around the world. In both monthly series, it was the 2008 crisis which raised them to their highest levels in the whole sample. Also, some peaks in these series occur close or inside the recessions periods shown in the light grey shaded area, though it is more evident in Figure 6b.

Table 4 - Monthly Measures of Idiosyncratic Variance – Brazilian Sample

Panel A shows the annualized mean and standard deviation of the monthly CSV measure proposed by Garcia, Mantilla-García, and Martellini (2014) and other competitive measures of idiosyncratic variance. In summary, four alternative measures were employed: the one based on the residuals of a multifactor regression according to the Sharpe (1964), Fama and French (1993, 2015), and Carhart (1997) models – CAPM, FF, FFC and FF5, respectively; the one based on gains from diversification available in Bali, Cakici, and Levy (2008) - Bali; the total variance, which is mostly idiosyncratic, given in Goyal and Santa-Clara (2003) - GS; and the difference between Goyal and Santa-Clara (2003) measure and the market variance - GS-M. Panel B reports the correlation matrix between the different measures. The sample period is from January 2000 to June 2016.

Panel A: Mean and Standard Deviation of Different Measures of Idiosyncratic Variance

	CSV^{CW}	$CAPM^{CW}$	FF^{CW}	FFC^{CW}	$FF5^{CW}$	$Bali^{CW}$	GS^{CW}	$GS-M^{CW}$	CSV^{EW}	$CAPM^{EW}$	FF^{EW}	FFC^{EW}	$FF5^{EW}$	$Bali^{EW}$	GS^{EW}	$GS-M^{EW}$
Mean	0,107	0,093	0,074	0,067	0,061	0,088	0,169	0,108	0,436	0,420	0,361	0,330	0,308	0,210	0,390	0,338
Standard deviaton	0,020	0,017	0,014	0,013	0,012	0,015	0,052	0,023	0,084	0,081	0,069	0,066	0,060	0,028	0,084	0,066

Panel B: Correlation Matrix between Different Measures of Idiosyncratic Variance

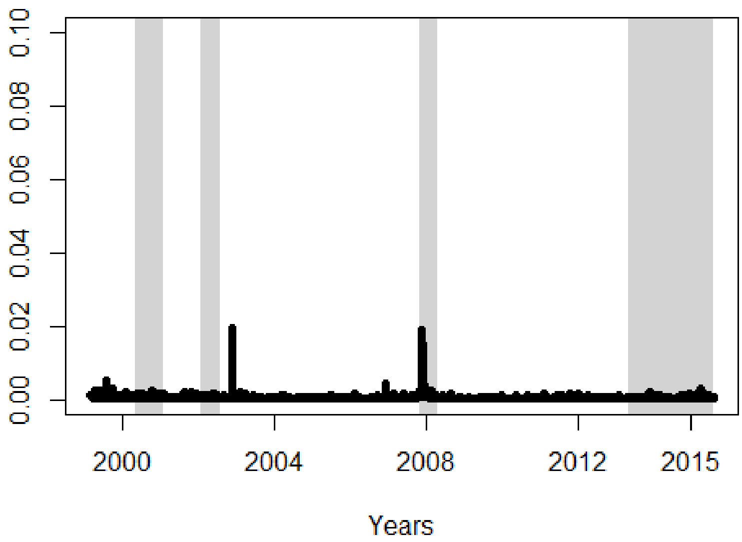
	CSV^{CW}	$CAPM^{CW}$	FF^{CW}	FFC^{CW}	$FF5^{CW}$	$Bali^{CW}$	GS^{CW}	$GS-M^{CW}$	CSV^{EW}	$CAPM^{EW}$	FF^{EW}	FFC^{EW}	$FF5^{EW}$	$Bali^{EW}$	GS^{EW}	$GS-M^{EW}$
CSV^{CW}	100	98,34	96,26	95,72	93,89	98,55	93,19	98,37	47,99	47,32	48,23	48,54	47,91	86,81	71,85	59,23
$CAPM^{CW}$		100	98,74	98,33	96,94	97,80	94,16	97,27	49,42	48,88	49,91	50,29	49,51	88,14	74,12	60,73
FF^{CW}			100	99,74	98,66	95,12	94,26	95,52	46,38	45,51	46,94	47,58	46,82	86,77	73,83	60,14
FFC^{CW}				100	98,68	94,48	94,32	95,26	46,02	45,15	46,59	47,35	46,56	86,36	73,48	59,69
$FF5^{CW}$					100	92,12	95,44	93,54	45,66	44,54	46,48	47,21	46,86	86,19	75,97	62,03
$Bali^{CW}$						100	90,62	97,08	47,12	46,47	46,95	47,26	46,35	85,77	68,81	56,34
GS^{CW}							100	93,70	46,73	45,87	47,58	48,08	48,03	85,29	78,29	62,06
$GS-M^{CW}$								100	47,67	46,92	47,82	47,94	47,49	86,56	73,11	60,69
CSV^{EW}									100	99,71	98,45	98,55	98,07	71,75	78,75	82,38
$CAPM^{EW}$										100	98,99	98,91	98,38	70,82	77,27	80,64
FF^{EW}											100	99,61	98,97	71,58	78,61	81,67
FFC^{EW}												100	98,99	71,73	78,91	81,79
$FF5^{EW}$													100	71,08	77,75	80,27
$Bali^{EW}$														100	87,62	80,18
GS^{EW}															100	96,62
$GS-M^{EW}$																100

Source: Author (2017)

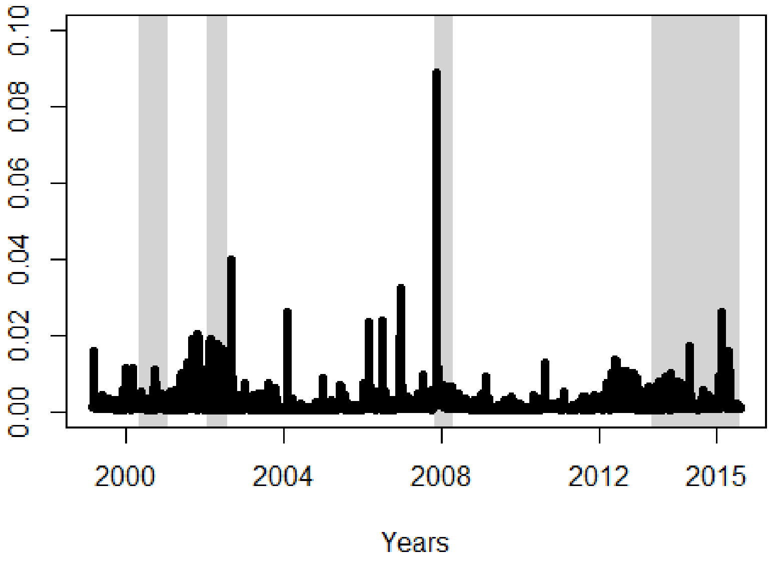
Figure 5 - Daily Idiosyncratic Variance – Brazilian Sample

The figure displays the daily idiosyncratic variance for the Brazilian stock market where a) represents the value-weighted measure and b) the equal-weighted version. The sample period is from January 2000 to June 2016.

a)



b)

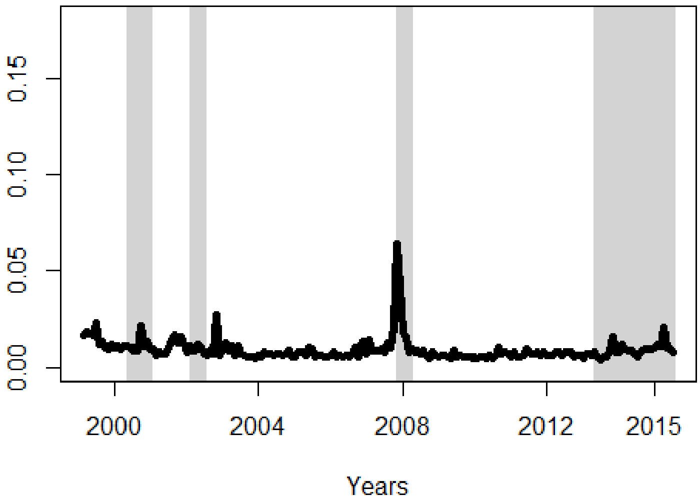


Source: Author (2017)

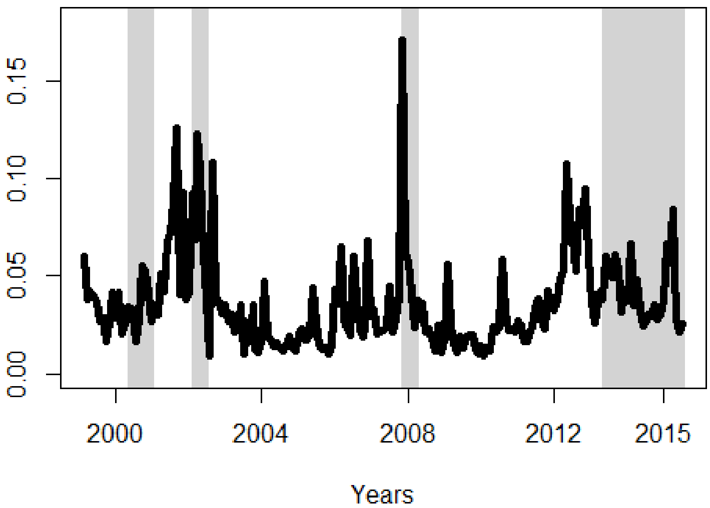
Figure 6 - Monthly Idiosyncratic Variance – Brazilian Sample

The figure displays the monthly idiosyncratic variance for the Brazilian stock market where a) represents the value-weighted measure and b) the equal-weighted version. The sample period is from January 2000 to June 2016.

a)



b)



Source: Author (2017)

4.1.2 Aggregate Idiosyncratic Variance over Time (Trend)

First, this subsection has no intention of providing a formal investigation of whether there is a trend in CSV and what may have caused it. In this sense, other studies as in Campbell et al. (2001), Xu and Malkiel (2003), Brandt et al. (2010), Fink et al. (2010) have studied if there is a trend in idiosyncratic risk and possible explanations to understand this phenomenon. For instance, Cao, Simin, and Zhao (2008) studied this research problem from a growth options perspective whereas Fink et al. (2010) focused on the impact of the firm maturity in understanding the increase in idiosyncratic risk during the Internet boom.

The purpose pursued here was to understand its behavior along the years and give a sense of whether it was necessary to include a detrended version of the cross-sectional variance (CSV) because, as pointed by Guo and Savickas (2006), if the measure has a trend, then results may be distorted by it. Table 5 presents the results of daily and monthly regressions of aggregate idiosyncratic variance over time during four subperiods and using the United States stock market.

Table 5 - Daily and Monthly Trend in CSV – United States Sample

Panel 1 shows the results of daily regressions of cross-sectional variance on time using four sample periods. The estimated regression is: $CSV_t = a + \beta T_t + \varepsilon_t$, where CSV_t represents either the equal or value-weighted idiosyncratic variance, α is the constant, β is the slope of the regressions, T_t is the trend, and ε_t is the model's residual. Panel 1A exhibits the equal-weighted weighting scheme and Panel 1B the value-weighted one. Panel 2 shows the monthly regressions counterparts. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from July 1963 to December 2014.

Panel 1: Daily				
	1963:7 - 1999:12	1963:7 - 2001:12	1963:7 - 2006:12	1963:7 - 2014:12
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 1A: Trend				
Constant	1,74E-04	1,13E-04	5,05E-04	8,68E-04
NW t - statistics	(3,154)	(1,874)	(7,862)	(12,747)
Coefficient	3,12E-07	3,31E-07	2,19E-07	1,30E-07
NW t - statistics	(21,219)	(21,713)	(12,396)	(8,097)
$Adj. R^2$	47,6%	49,4%	26,9%	7,8%

Continued

	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 1B: Trend				
Constant	1,83E-04	1,13E-04	1,82E-04	2,41E-04
NW t - statistics	(14,748)	(5,190)	(10,518)	(16,453)
Coefficient	3,05E-08	6,46E-09	3,29E-08	1,85E-08
NW t - statistics	(9,881)	(81,572)	(6,682)	(5,156)
$Adj. R^2$	13,0%	17,6%	9,3%	36,3%
Panel 2: Monthly				
	1963:7 - 1999:12	1963:7 - 2001:12	1963:7 - 2006:12	1963:7 - 2014:12
	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 2A: Trend				
Constant	0,003	0,002	0,010	1,81E-02
NW t - statistics	(1,238)	(0,756)	(2,974)	(5,079)
Coefficient	2,00E-04	1,00E-04	9,68E-05	5,74E-05
NW t - statistics	(10,293)	(11,706)	(4,811)	(3,349)
$Adj. R^2$	60,4%	62,6%	36,2%	14,0%
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 2B: Trend				
Constant	0,004	0,002	0,004	0,0050
NW t - statistics	(6,352)	(2,093)	(4,138)	(6,819)
Coefficient	1,37E-05	2,34E-05	1,00E-05	8,23E-06
NW t - statistics	(4,589)	(3,385)	(2,626)	(2,107)
$Adj. R^2$	25,1%	27,5%	14,5%	5,6%

Source: Author (2017)

The trend analysis strongly suggests that CSV has an upward movement in both daily and monthly frequencies and is independent from the weighting scheme employed. This result is in accordance with Campbell *et al.* (2001), using monthly frequency and taking into account idiosyncratic risk at the firm-level. Still, Brandt *et al.* (2010) point out that there was no upward trend, but rather an episodic phenomenon linked to low-priced stocks with high retail trading.

For the four sample periods, it was found that the variable time has a positive and statistically coefficient showing that idiosyncratic variance has increased over time. In line with our results, Guo and Savickas (2006) also do not discuss what may have contributed to a trend in the data, but the authors report their analysis of aggregate idiosyncratic variance based on Goyal and Santa-Clara's (2003) approach using a detrended version of idiosyncratic risk or a trend variable trying to account for the positive significant trend coefficient revealed through their analysis via quarterly data. Therefore, the main regressions verifying if idiosyncratic variance forecasts market returns were also tested including a detrended version of the measure.

The same regressions were run to the CSV constructed from Brazilian stocks to verify if a detrended measure should be estimated and to try to reach a better understanding of idiosyncratic variance and avoid any spurious regressions (GUO & SAVICKAS, 2006).

Table 6 - Daily and Monthly Trend – Brazilian Sample

Panel 1 shows the results of daily regressions of cross-sectional variance on time using three sample periods. The estimated regression is: $CSV_t = \alpha + \beta T_t + \varepsilon_t$, where CSV_t represents either the equal or value-weighted idiosyncratic variance, α is the constant, β is the slope of the regressions, T_t is the trend, and ε_t is the model's residual. Panel 1A exhibits the equal-weighted weighting scheme and Panel 1B the value-weighted measure. Panel 2 shows the monthly regressions. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000 - 2016	2000-2007	2008-2016
Panel 1: Daily			
	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 1A: Trend			
Constant	0,002	0,002	0,001
NW t - statistics	(13,571)	(13,383)	(6,353)
Coefficient	1,16E-08	-5,52E-07	4,40E-07
NW t - statistics	(0,234)	(-4,061)	(2,844)
$Adj.R^2$	0,0%	1,3%	1,0%

Continued

	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 1B: Trend			
Constant	5,23E-04	6,33E-04	5,21E-04
NW t - statistics	(20,968)	(19,963)	(6,728)
Coefficient	-4,46E-08	-1,91E-07	-9,70E-8
NW t - statistics	(-4,968)	(-8,115)	(-1,797)
$Adj.R^2$	0,8%	3,9%	0,9%
	2000 - 2016	2000-2007	2008-2016
Panel 2: Monthly			
	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 2A: Trend			
Constant	0,036	0,047	0,027
NW t - statistics	(6,112)	(6,5799)	(3,603)
Coefficient	4,07E-06	-2,37E-04	1,90E-04
NW t - statistics	(0,085)	(-2,355)	(1,624)
$Adj.R^2$	-0,5%	5,9%	4,4%
	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 2B: Trend			
Constant	0,011	0,013	0,011
NW t - statistics	(10,848)	(11,917)	(3,245)
Coefficient	-1,93E-05	-8,21E-05	-4,15E-05
NW t - statistics	(-2,515)	(-4,545)	(-0,877)
$Adj.R^2$	3,0%	29,8%	2,0%

Source: Author (2017)

Daily CSV^{EW} demonstrates a downward trend in the first period of the sample (from January 2000 to December 2007) followed by an upward movement in the second half of the sample (January 2008 to June 2016), but when the complete timespan is considered there is no evidence of a significant trend. In its monthly version, the same behavior is observed with the exception that in the second half it is non-significant.

In contrast, CSV^{CW} has a donward trend in all the three periods analyzed being marginally significant (non-significant) only in the second half of the sample in the daily (monthly) data. Again, it becomes clear that when disaggregating by the weighting scheme both measures have a different pattern, where the equal-weighted measure reflects the stocks of small firms and value-weighted that of big firms.

Angelidis (2010) utilizes a sample of countries which are part of the Morgan Stanley Capital International (MSCI) Emerging Markets, including Brazil, and computes the aggregate idiosyncratic volatility taking the standard deviation of residuals relative to the CAPM model and a value-weighting scheme across the stocks in the sample. The author reports a negative though insignificant trend using Brazilian stocks for the period from December 1994 to May 2007.

Also, Costa, Mazzeu, and Costa Junior (2016) point out a negative though significant trend at the firm-level idiosyncratic volatility following the procedure outlined in Campbell et al. (2001) in the Brazilian case. Then, the results here corroborate that there is a small negative trend in idiosyncratic variance.

4.1.3 Daily and Monthly Predictive Regressions using Aggregate Idiosyncratic Variance

This subsection shows the analysis of whether idiosyncratic variance is important to predict one-day/month ahead market returns using both equal- and value-weighted market returns. Table 7 exhibits the role of idiosyncratic variance in forecasting market returns. The sample starts in July 1963 as in Garcia, Mantilla-García, and Martellini (2014), since they made it comparable to the sample period in Goyal and Santa-Clara (2003), Bali *et al.* (2005), and Wei and Zhang (2005).

These estimated regressions are useful in verifying if aggregate idiosyncratic variance predicts market returns using a daily frequency of data. This is new evidence regarding the relation between idiosyncratic risk and return, as mentioned by Garcia, Mantilla-García, and Martellini (2014), since the main empirical works are based on a monthly frequency. Also, it can be useful for those active traders who constantly rebalance their portfolios and not only to those traders or managers who adopt a long investment horizon.

Table 7 shows that $CSV^{EW}$and CSV^{CW} are both important to predict one-day-ahead market returns and present a positive signal throughtout the sample period. It is shown that the coefficients regarding daily predictability of market returns by aggregate idiosyncratic variance until 2006 are close to those reported by Garcia, Mantilla-García, and Martellini (2014). Also, updating the time span to 2014, both equal- and value-weighted CSV interpretations remain unchanged. This means that by looking at the cross-sectional variance of stocks as a measure of idiosyncratic risk has been robust in the sample across time and, more importantly, it matters in a daily frequency and that this portion of the risk should be taken into account by investors.

Table 7 - Daily Predictability of Market Returns by CSV – United States Sample

Panel A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance for four sample periods. Panel B shows the value-weighted daily market excess returns on the daily lagged value-weighted cross-sectional variance. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal- or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from July 1963 to December 2014.

	1963:7 - 1999:12	1963:7 - 2001:12	1963:7 - 2006:12	1963:7 - 2014:12
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel A: Forecasting r^{EW}				
Constant	-3,57E-04	-3,34E-04	-1,65E-04	5,30E-05
NW t - statistics	(-1,619)	(-1,568)	(-0,829)	(0,234)
Coefficient	0,589	0,534	0,453	0,317
NW t - statistics	(5,164)	(5,032)	(4,419)	(2,744)
$Adj.R^2$	1,1%	1,0%	0,7%	0,4%
	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel B: Forecasting r^{CW}				
Constant	-8,58E-04	-2,50E-04	-2,16E-04	-1,25E-04
NW t - statistics	(-3,983)	(-0,978)	(-0,905)	(-0,062)
Coefficient	3,442	1,254	1,201	1,005
NW t - statistics	(6,266)	(2,046)	(2,049)	(2,025)
$Adj.R^2$	0,9%	0,2%	0,2%	0,1%

Source: Author (2017)

These results are now expanded to a monthly frequency to verify the ability of aggregate idiosyncratic variance to forecast market returns, as ginve in Table 8.

Table 8 - Monthly Predictability of Market Returns by CSV – United States Sample

Panel A shows the results of a one-month-ahead predictive regression of the value-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance for four sample periods. Panel B shows the results changing only the market returns to its equal-weighted version. Panel C shows the results for the regression of value-weighted monthly market excess returns on the monthly lagged value-weighted cross-sectional variance. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from July 1963 to December 2014.

	1963:7 - 1999:12	1963:7 - 2001:12	1963:7 - 2006:12	1963:7 - 2014:12
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel A: Forecasting r^{CW}				
Constant	-0,003	-9,11E-05	0,001	0,003
NW t - statistics	(-0,780)	(-0,025)	(0,377)	(0,962)
Coefficient	0,255	0,133	0,097	0,060
NW t - statistics	(2,734)	(1,579)	(1,232)	(0,754)
$Adj. R^2$	1,5%	0,3%	0,1%	0,0%
Panel B: Forecasting r^{EW}				
Constant	-0,002	-0,002	-1,85E-04	-2,70E-04
NW t - statistics	(-0,336)	(-0,345)	(-0,037)	(-0,061)
Coefficient	0,283	0,257	0,231	0,228
NW t - statistics	(2,564)	(2,559)	(2,488)	(2,284)
$Adj. R^2$	1,1%	1,0%	0,8%	1,0%
	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel C: Forecasting r^{CW}				
Constant	0,002	0,008	0,009	0,011
NW t - statistics	(0,407)	(2,422)	(2,706)	(4,350)
Coefficient	0,552	-0,462	-0,500	-0,829
NW t - statistics	(0,880)	(-1,317)	(-1,505)	(-2,468)
$Adj. R^2$	-0,1%	0,2%	0,2%	1,1%

Source: Author (2017)

Results shown in Table 8 verify the predictability power of idiosyncratic variance in a monthly frequency. In the first, second and third periods in Panel A and B, there is evidence of a positive relation between aggregate idiosyncratic variance and one-month-ahead market returns, though only statistically significant for all three periods when using CSV^{EW} to predict equal-weighted market returns (r^{EW}), but with a loss in the statistical power over time. Additionally, as pointed out by Garcia, Mantilla-García, and Martellini (2014), for Panel A, in their three analyzed periods and confirmed here in the extended sample, there is support for the evidence shown in Bali *et al.* (2005) and Wei and Zhang (2005) that equal-weighted idiosyncratic variance cannot predict value-weighted market returns after including more data to the analysis.

In contrast, when using CSV^{CW} in Panel C, it shows that increasing the timespan until 2014, value-weighted CSV became statistically significant, i.e. it helps to predict one-month-ahead value-weighted market returns (r^{CW}), as opposed to the outlined by traditional asset pricing theory, but supporting the idea that if investors do not diversify away their portfolios, then idiosyncratic variance may matter. More interestingly, the relation has a negative sign in opposition to the positive one observed in the equal weighting scheme. Overall, the results until 2006 are similar to those documented by Garcia, Mantilla-García, and Martellini (2014).

Table 9 describes the daily predictive ability of idiosyncratic variance for market returns in Brazilian case.

Table 9 - Daily Predictability of Market Returns by CSV – Brazilian Sample

Panel A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance for three sample periods. Panel B shows the value-weighted daily market excess returns on the daily lagged value-weighted cross-sectional variance. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000 - 2016	2000-2007	2008-2016
	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel A: Forecasting r^{EW}			
Constant	0,001	0,002	0,001
NW t - statistics	(3,252)	(3,631)	(1,573)
Coefficient	-0,296	-0,231	-0,356
NW t - statistics	(-1,762)	(-1,546)	(-1,368)
$Adj. R^2$	0,3%	0,1%	0,5%
	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel B: Forecasting r^{CW}			
Constant	2,72E-04	4,62E-04	7,26E-05
NW t - statistics	(0,430)	(1,042)	(0,074)
Coefficient	-0,334	-0,070	-0,544
NW t - statistics	(-0,228)	(-0,104)	(-0,226)
$Adj. R^2$	0,0%	-0,1%	0,0%

Source: Author (2017)

With Brazilian data, the overall findings for daily CSV do not support those for the United States by reporting that aggregate idiosyncratic variance does not have an association to market returns for equal- and value-weighted schemes. The only exception is $\mathrm{CSV}^{\mathrm{EW}}$, which appears to be marginally significant and negative in the whole period. Next, the analysis is turned to monthly data.

Table 10 - Monthly Predictability of Market Returns by CSV – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the value-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance for three sample periods. Panel B shows the results changing only the market returns to its equal-weighted version. Panel C shows the results for the regression of value-weighted monthly market excess returns on the monthly lagged value-weighted cross-sectional variance. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to December 2016.

	2000-2016	2000-2007	2008-2016
	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel A: Forecasting r^{CW}			
Constant	0,004	0,017	-0,007
NW t - statistics	(0,531)	(1,475)	(-0,699)
Coefficient	-0,092	-0,276	0,089
NW t - statistics	(-0,599)	(-1,023)	(0,359)
$Adj. R^2$	-0,4%	0,1%	-0,9%
Panel B: Forecasting r^{EW}			
Constant	0,008	0,021	-0,003
NW t - statistics	(0,771)	(1,620)	(-0,229)
Coefficient	-0,042	-0,131	0,046
NW t - statistics	(-0,189)	(-0,442)	(0,194)
$Adj. R^2$	-0,5%	-0,9%	-1,0%
	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel C: Forecasting r^{CW}			
Constant	0,004	0,023	-0,003
NW t - statistics	(0,742)	(1,392)	(-0,338)
Coefficient	-0,392	-1,686	-0,098
NW t - statistics	(-0,785)	(-1,028)	(-0,012)
$Adj. R^2$	-0,4%	0,1%	-1,0%

Source: Author (2017)

Perhaps not surprisingly, aggregate idiosyncratic variance measured by CSV does not reveal itself to play a role in predicting Brazilian market returns. In seven out of nine predictive regressions, its coefficient is negative, while it is insignificant and positive in the two remaining regressions although also indistinguishable from zero. According to these results previously reported, Hypothesis 1 which states that CSV plays a central role in predicting market returns was only verified for the United States sample in a daily and monthly frequency using both equal – and value-weighted market returns and considering the whole sample.

In addition, these analysis document that idiosyncratic variance from an emerging country perspective does not support Merton's (1987) theory that when investors do not diversify away their portfolios idiosyncratic risk should matter. As in Guo and Savickas (2006), for the cases where a significant coefficient associated to the variable time was found (trend analysis presented in section 4.1.2), then a detrended CSV measure was utilized. This detrended version corresponded to the residuals of the regression of CSV on time. The results were virtually unchanged and were not tabulated to conserve space.

Given these results for the Brazilian case, three new attempts to understand whether aggregate idiosyncratic variance is relevant were made in the next analysis. First, daily and monthly observations were split into up and down markets according to the return of the Ibovespa stock index return minus the return of the risk-free rate (namely, the CDI) or in periods of expansion and contraction given by the sign of the real per capita Gross Domestic Product (GDP) change as well adding dividend yield and expected and unexpected market illiquidity. The results are summarized in Tables 11 and 12.

Table 11 shows the results of the predictability power of idiosyncratic variance for market returns in four different specifications because when the previous results are considered, where it was observed independently of the market state, a non-significant predictability power was reported.

Surprisingly, considering only the daily frequency (Panel 1), CSV^{EW} is non-significant only in the up-market markets and, differently from the previous results, it now has statistical power in down-market, expansion and contraction periods. In opposition, CSV^{CW} has forecasting ability for market returns only in up and down markets. Changing the attention to the monthly frequency seems to not make a considerable effect as there are only two cases for which it is marginally significant: CSV^{EW} in up-market to predict value-weighted market returns and CSV^{EW} in expansion period to predict equal-weighted market returns.

Table 11 - Predictability of Market Returns by CSV Conditioning on the State of the Market – Brazilian Sample

Panel 1A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance in up/down markets given by the return of the Ibovespa minus the risk-free rate and in expansion/contraction periods given by the sign of real per capita GDP change. Panel 1B shows its value-weighted daily counterpart. Panels 2A (Panel 2B) provide the results for one-month-ahead predictive regression of the value-weighted (equal-weighted) monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance according to the state of the market. Panel 2C comprehends the value-weighted cross-sectional variance counterpart. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

Panel 1: Daily	Up	Down	Expansion	Contraction
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 1A: Forecasting r^{EW}				
Constant	0,009	-0,007	0,001	0,001
NW t - statistics	(30,101)	(-20,288)	(3,544)	(0,979)
Coefficient	0,187	-0,579	-0,517	0,256
NW t - statistics	(1,415)	(-4,865)	(-3,308)	(1,989)
$Adj. R^2$	0,2%	2,4%	1,0%	0,2%
	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 1B: Forecasting r^{CW}				
Constant	0,008	-0,009	1,90E-04	4,00E-04
NW t - statistics	(15,143)	(-14,970)	(0,214)	(0,936)
Coefficient	5,464	-3,426	-0,584	-0,105
NW t - statistics	(4,466)	(-2,347)	(-0,273)	(-0,142)
$Adj. R^2$	5,3%	4,0%	0,0%	-0,1%

Continued

Panel 2: Monthly	Up	Down	Expansion	Contraction
	CSV^{EW}	CSV^{EW}	CSV^{EW}	CSV^{EW}
Panel 2A: Forecasting r^{CW}				
Constant	0,061	-0,042	0,006	0,007
NW t - statistics	(8,095)	(-5,676)	(0,642)	(0,419)
Coefficient	-0,314	-0,045	-0,263	0,020
NW t - statistics	(-1,842)	(-0,257)	(-1,233)	(0,052)
$Adj.R^2$	2,5%	0,0%	0,4%	-1,8%
Panel 2B: Forecasting r^{EW}				
Constant	0,062	-0,035	0,014	-0,003
NW t - statistics	(6,799)	(0,012)	(1,275)	(-0,127)
Coefficient	-0,108	-0,145	-0,358	0,409
NW t - statistics	(-0,529)	(0,562)	(-1,679)	(1,075)
$Adj.R^2$	0,0%	0,0%	1,2%	-0,4%
	CSV^{CW}	CSV^{CW}	CSV^{CW}	CSV^{CW}
Panel 2C: Forecasting r^{CW}				
Constant	0,049	-0,040	0,005	-0,005
NW t - statistics	(5,903)	(-5,475)	(0,594)	(-0,237)
Coefficient	0,050	-0,488	-0,886	1,456
NW t - statistics	(0,061)	(-0,749)	(-1,425)	(0,667)
$Adj.R^2$	0,0%	0,0%	0,2%	-0,9%

Source: Author (2017)

Next, Table 12 brings the analysis of aggregate idiosyncratic variance when controlling for aggregate dividend yield and unexpected and expected illiquidity on a monthly frequency. The intention was to verify how the cross-sectional variance acts when jointly considered with measures which are considered as important predictors in the literature. Then, three models specifications were run: Model 1 uses only divided yield to predict market returns; Model 2 adds either equal- or value-weighted aggregate idiosyncratic variance; Model 3 includes the liquidity measures to the previous model.

Table 12 - Monthly Predictability of Market Returns by CSV, Aggregate Dividend Yield and Market Liquidity – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance, aggregate dividend yield and unexpected and expected illiquidity measures based on Amihud (2002). Panel B shows the value-weighted monthly market excess returns on the monthly lagged value-weighted cross-sectional variance, aggregate dividend yield and unexpected and expected illiquidity measures based on Amihud (2002). Model 1 uses only dividend yield as a predictor variable. Model 2 uses equal or value- weighted aggregate idiosyncratic risk and dividend yield jointly. Model 3 adds the liquidity measures to the previous model. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from February 2001 to June 2016.

Panel A: Forecasting r^{EW}	Constant	CSV^{EW}	DY	$Illq^U$	$Illq^E$	$Adj.R^2$
1	-0,048		2,111			4,7%
t-stat	(-2,578)		(3,240)			
2	-0,047	-0,019	2,108			4,2%
t-stat	(-2,500)	(-0,089)	(3,263)			
3	0,036	-0,159	2,278	0,003	0,007	4,0%
t-stat	(0,470)	(-0,708)	(3,507)	(0,766)	(1,085)	
Panel B: Forecasting r^{CW}	Constant	CSV^{CW}	DY	$Illq^U$	$Illq^E$	$Adj.R^2$
1	-0,042		1,720			3,9%
t-stat	(-2,671)		(2,900)			
2	-0,040	-0,377	1,743			3,5%
t-stat	(-2,368)	(-0,496)	(2,921)			
3	0,214	-0,900	1,842	0,012	0,016	5,5%
t-stat	(1,541)	(-1,147)	(3,112)	(1,622)	(1,828)	

Source: Author (2017)

The main investigation relies upon the role of aggregate risk when jointly employed with other control variables to predict market returns. In this sense, Table 12 reports that both CSV^{EW} and CSV^{CW} do not predict one-month-ahead market returns in any of the models reviewed here whereas, in contrast, DY seems to be a pervasive variable. The unexpected and expected illiquidity measures, $Illq^U$ and $Illq^{\mathrm{E}}$, respectively, are not capable of predicting market returns.

In summary, the daily analysis present that conditioning on the market state there are limited evidences that idiosyncratic risk can be used to predict market returns, although the overall results do not corroborate that idiosyncratic variance is relevant in the Brazilian stock market. Interestingly, Garcia, Mantilla-García, and Martellini (2014) use the relative T-bill rate, aggregate dividend yield, and term spread as macroeconomic control variables jointly with cross-sectional variance in a daily frenquency as well as the default credit spread on a monthly basis.

Their results show that, except when using the value-weighted aggregate idiosyncratic risk to predict value-weighted market returns in a monthly frequency, equal- and value-weighted idiosyncratic variance remain statistically significant and, specifically concerning the dividend yield, the variable is non-significant. These results cannot be directly compared to those discussed here as they make use of different control variables, but the notion that opposite results are shown in the Brazilian case at least provides an indication between the differences in the markets.

4.1.4 Predictability of Market Returns using Aggregate Idiosyncratic Variance and Market Variance

The analysis is turned to the investigation of whether cross-sectional variance and market variance together can predict one-day/month ahead market returns using the United States stocks, as given in Table 13.

The investigation of whether controlling for the market variance together with the cross-sectional variance helps to predict market returns is very important. Guo and Savickas (2006) affirm that the positive tradeoff between return-risk could not be recovered in some studies due to an omitted factor problem – the idiosyncratic volatility was not controlled for. In this sense, they found that when testing a predictive regression of market excess return by both idiosyncratic volatility and market variance, the former is significantly and negatively related and the latter shows a significantly and positively association with market returns.

In this specific case, when equal-weighted market excess returns are considered and using both idiosyncratic variance and market variance, CSV^{EW} is highly significant while MV^{EW} is only marginally significant in daily frequency and until December 2006. In addition, CSV^{CW} and MV^{CW}, when taken into account separately, prove to be important in daily data in the first part of the sample, but lose power in a bivariate regression and both are not significant in the whole sample.

Table 13 - Daily and Monthly Predictability of Market Returns by CSV and Market Variance – United States Sample

Panel 1A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance and/or equal-weighted market variance. Panel 1B shows the value-weighted daily market excess returns on the daily lagged value-weighted cross-sectional variance and/or value-weighted market variance. Panel 2A and 2B shows the corresponding counterparts from monthly data. The estimated regression is: $r_{t+1} = a + \beta v_t + \gamma m v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market excess returns in period t+1, a is the constant of the model, β is the slope coefficient related to the aggregate idiosyncratic risk; v_t represents either the equal or value-weighted idiosyncratic variance in period t, γ is the slope coefficient related to the market variance; mv_t is either the equal or value-weighted market variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from July 1963 to December 2014.

		1963:7 - 2006:12				1963:7 - 2014:12			
Panel 1: Daily									
Panel 1A: Forecasting r^{EW}		Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
	1	-1,65E-04	0,453		0,7%	5,30E-05	0,317		0,4%
	t-stat	(-0,829)	(4,419)			(0,234)	(2,744)		
	2	0,001		-1,673	0,0%	6,13E-04		-0,285	0,0%
	t-stat	(4,825)		(-0,765)		(4,399)		(-0,148)	
	3	-6,1E-05	0,504	-3,817	0,8%	1,20E-04	0,360	-2,168	0,5%
	t-stat	(-0,297)	(5,130)	(-1,860)		(0,542)	(2,481)	(-1,022)	
Panel 1B: Forecasting r^{CW}		Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
	1	-2,16E-04	1,201		0,2%	-1,25E-04	1,005		0,1%
	t-stat	(-0,905)	(2,049)			(-0,062)	(2,025)		
	2	5,58E-06		2,825	0,1%	1,32E-04		1,162	0,0%
	t-stat	(0,048)		(2,117)		(1,040)		(0,861)	
	3	-2,23E-04	1,148	0,352	0,2%	-1,17E-04	1,131	-0,593	0,1%
	t-stat	(-1,090)	(1,352)	(0,151)		(-0,593)	(1,612)	(-0,316)	
Panel 2: Monthly									
Panel 2A: Forecasting r^{EW}		Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
	1	-1,85E-04	0,231		0,8%	-2,70E-04	0,228		1,0%
	t-stat	(-0,037)	(2,488)			(-0,061)	(2,284)		
	2	0,008		-0,216	-0,2%	0,009		-0,819	0,1%
	t-stat	(2,870)		(-0,166)		(3,569)		(-0,839)	
	3	6,52E-05	0,251	-0,991	0,7%	-0,002	0,351	-2,177	2,1%
	t-stat	(0,013)	(2,577)	(-0,845)		(-0,329)	(2,868)	(-2,472)	
Panel 2B: Forecasting r^{CW}		Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
	1	0,009	-0,500		0,2%	0,011	-0,829		1,1%
	t-stat	(2,706)	(-1,505)			(4,350)	(-2,468)		
	2	0,006		-0,641	0,0%	0,007		-1,071	0,9%
	t-stat	(2,633)		(-0,952)		(4,108)		(-2,642)	
	3	0,009	-0,488	-0,031	0,0%	0,011	-0,606	-0,480	1,1%
	t-stat	(2,641)	(-1,153)	(-0,036)		(4,190)	(-1,755)	(-0,967)	

Source: Author (2017)

In the extended period in a monthly frequency, CSV^{EW} maintains its relevance to forecast equal-weighted market returns even in the presence of MV^{EW}, which in turn is also significant when employed together. In contrast, when using the value-weighted scheme, CSV^{CW} and MV^{CW} were only important when used separately, in other words, the value-weighted measure of firm-specific risk was not robust to the inclusion of the value-weighted market variance at the 5% level. Overall, the results are similar to those documented by Garcia, Mantilla-García, and, Martellini (2014) in a period lasting until 2006.

Proceeding further, the same regressions with cross-sectional variance and market variance using equal- and value-weighted schemes to predict equal- and value-weighted market returns were run using the Brazilian stock market sample to verify if there was a pattern to the data. These results are shown in Table 14 using both daily and monthly frequency.

Table 14 - Daily and Monthly Predictability of Market Returns by CSV and Market Variance – Brazilian Sample

Panel 1A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance and/or equal-weighted market variance. Panel 1B shows the value-weighted daily market excess returns on the daily lagged value-weighted cross-sectional variance and/or value-weighted market variance. Panel 2A and 2B shows the corresponding counterparts from monthly data. The estimated regression is: $r_{t+1} = a + \beta v_t + \gamma m v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market excess returns in period t+1, a is the constant of the model, β is the slope coefficient related to the aggregate idiosyncratic risk, v_t represents either the equal or value-weighted idiosyncratic variance in period *t*, γ is the slope coefficient related to the market variance, mv_t is either the equal or value-weighted market variance in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000 - 2016			
Panel 1: Daily				
Panel 1A: Forecasting r^{EW}	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
1	0,001	-0,296		0,3%
t-stat	(3,252)	(-1,762)		
2	0,001		0,610	0,0%
t-stat	(0,996)		(0,202)	
3	0,001	-0,312	1,396	0,3%
t-stat	(1,655)	(-1,793)	(0,454)	

Continued

Panel 1B: Forecasting r^{CW}	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
1	2,72E-04	-0,334		0,0%
t-stat	(0,430)	(-0,228)		
2	-4,75E-04		2,650	0,1%
t-stat	(-0,877)		(1,055)	
3	-3,12E-04	-0,863	3,571	0,2%
t-stat	(-0,469)	(-0,568)	(1,333)	
		2000 - 2016		
Panel 2: Monthly				
Panel 2A: Forecasting r^{EW}	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
1	0,008	-0,042		-0,5%
t-stat	(0,771)	(-0,189)		
2	0,010		-0,821	-0,1%
t-stat	(1,457)		(-0,969)	
3	0,008	0,049	-0,919	-0,6%
t-stat	(0,787)	(0,175)	(-0,857)	
Panel 2B: Forecasting r^{CW}	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
1	0,004	-0,392		-0,4%
t-stat	(0,742)	(-0,785)		
2	0,004		-0,699	0,1%
t-stat	(0,868)		(-1,070)	
3	-0,003	1,321	-1,641	0,0%
t-stat	(-0,275)	(0,877)	(-1,317)	

Source: Author (2017)

Previous regressions show that, for the Brazilian case, aggregate idiosyncratic variance is not a robust and relevant variable to predict market returns alone. Table 14 adds the market variance to verify if it affects the relation. The results show that when market variance is included in the predictive regressions, it does not play a significant role in association to market returns neither in daily nor monthly frequency and only CSV^{EW} has marginal power to predict daily equal-weighted market returns. The adjusted R^2 displayed are very low ranging from -0,6% to 0,3% indicating the lack of power by the explanatory variables tested herein. In addition, as a robustness test, the largest values of the market variance measures were also replaced by their second largest one to control for the potential issue caused by extreme observation. The results were virtually unchanged and they were not tabulated here to conserve space.

The results in Table 13, using the United States stocks, show that except for the daily value-weighted measures in Panel 1B, all the remaining show a negative sign for market variance, though mostly without statistical power. For the Brazilian case, it is negative in all the regressions with monthly data, but still non-significant.

Guo and Savickas (2006) recover a positive market volatility signal when accounting for idiosyncratic risk and posit that previous authors fail to find this relation because of an omitted variable problem. Guo and Savickas (2008) reason that, in quarterly data, both aggregate idiosyncratic volatility and market variance are more relevant in predicting market returns probably due to the dependency of realized market volatility on a long-distributed past information of daily returns, as affirmed by Ghysels, Santa-Clara, and Valkanov (2005). Therefore, this negative effect may be caused by the nature of the cross-sectional variance which is based on daily returns.

4.1.5 Predictability of Market Returns using Aggregate Idiosyncratic Variance and Investor Sentiment

Now, the focus relies on the aggregate idiosyncratic variance's capability of predicting market excess returns controlling for the investor sentiment and market variance using the United States sample as shown in Table 15.

Table 15 - Predictability Regressions of Market Returns by CSV Controlling for Market Variance and Investor Sentiment – United States Sample

Panel A exhibits the monthly predicitive regressions of the equal-weighted market excess returns by lagged equal-weighted cross-sectional variance, lagged equal-weighted market variance and lagged investor sentiment. Panel B shows the value-weighted counterparts. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) correct t-statistics are shown in parentheses. The sample period is from July 1965 to December 2014.

Panel A: Equal-weighted measures

	Constant	CSV^{EW}	MV^{EW}	$SENT$	$Adj.R^2$
Forecasting r^{EW}	-0,005	0,435	-2,471	-0,009	4,5%
t-stat	(-0,963)	(3,782)	(-0,027)	(-3,973)	

Continued

Panel B: Cap-weighted measures					
	Constant	CSV^{CW}	MV^{CW}	$SENT$	$Adj.R^2$
Forecasting r^{CW}	0,010	-0,478	-0,006	-0,002	1,1%
t-stat	(3,112)	(-1,139)	(-1,028)	(-1,148)	

Source: Author (2017)

Table 15 tests the relation between aggregate idiosyncratic variance and returns when market variance and investor sentiment are included in the predictive regressions. As can be seen, CSV^{EW} retains its forecasting power even in the presence of MV^{EW} and the investor sentiment variable of Baker and Wurgler (2006), in which case the latter is also significant, though negatively.

In contrast, none of the variables has statistical significance to predict the next period's value-weighted market returns. The lack of power concerning CSV^{CW} is in a certain degree expected as the previously reported results available for the United States has shown that it is non-significant at the 5% level when adding the market variance to the model. Moreover, this negative relation of investor sentiment and market returns in both cases (and significant in the r^{EW}) was already expected to the United States market as found in Brown and Cliff (2005) and Schmeling (2009) and this means that when sentiment optimism is lower, then next period's market return is higher.

Next, the intention was also to evaluate whether and how market variance and investor sentiment affect the relation between idiosyncratic risk and market returns in Brazilian case, as can be seen on Table 16.

Table 16 - Predictability of Market Returns by CSV and Controlling for Market Variance and Investor Sentiment – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the following monthly lagged variables: equal-weighted cross-sectional variance, equal-weighted market variance, index of consumer confidence, index of economic activity, inflation and interest rate, respectively. Panel B shows the value-weighted counterparts. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2003 to June 2016.

Panel A: Equal-weighted measures								
	Constant	CSV^{EW}	MV^{EW}	ICC	$IBC-Br$	$IPCA$	CDI	$Adj.R^2$
Forecasting r^{EW}	0,375	0,422	-2,052	-1,00E-04	-0,002	-1,657	-5,181	10,4%
t-stat	(2,919)	(1,662)	(-2,201)	(-0,391)	(3,617)	(-1,101)	(-1,590)	
Panel B: Cap-weighted measures								
	Constant	CSV^{CW}	MV^{CW}	ICC	$IBC-Br$	$IPCA$	CDI	$Adj.R^2$
Forecasting r^{CW}	0,213	2,715	-2,785	-1,00E-04	-0,001	-1,833	-2,927	5,2%
t-stat	(2,163)	(1,148)	(-1,369)	(-0,441)	(-2,643)	(-1,526)	(-1,172)	

Source: Author (2017)

As mentioned before, investor sentiment in the Brazilian analysis is estimated by using the Consumer Confidence Index (Índice de Confiança do Consumidor - ICC) and adding fundamental variables to take into account the exposition to macroeconomic changes. Focusing on the equal-weighted market returns, it can be seen that the inclusion of proxies to account for investor sentiment contrasts to those previously reported results, which considered only $\mathrm{CSV^{EW}}$ and $\mathrm{MV^{EW}}$.

In the forecasting regressions of the equal-weighted market returns, it is found that $\mathrm{MV^{EW}}$ is significant at the 5% level and it has a negative sign. In contrast, aggregate idiosyncratic variance, investor sentiment and fundamental variables do not show to play a role for the performed regression, with exception of IBC-Br, which represents the index of economic activity. Looking at the forecasting for the value-weighted market returns, it is observed that neither aggregate idiosyncratic variance nor market variance and investor sentiment are relevant predictors, but IBC-Br remains statistically important.

In general, regarding Hypothesis 2 and taking into account the importance of the variables employed in the analysis with a 5% level of statistical significance, for the United States sample the null hypothesis that idiosyncratic variance does not matter in the presence of market variance and investor sentiment can be rejected considering only the equal-weighted measure while for the Brazilian case it is rejected in both weighting schemes.

Lastly, these results also provide evidence that investor sentiment, as measured by a consumer index, is not related to market returns, but has the expected negative sign.

4.1.6 Predictability of Market Returns using Higher Cross-Sectional Moments

The next step was to verify if higher cross-sectional idiosyncratic moments beyond variance, specifically the skewness and kurtosis, play a role in predicting market returns. Results are displayed in Table 17 for the United States sample. Cross-sectional skewness (CSS) may matter to investors as shown in Table 17. CSS is positively and significantly related to both equal- and value-weighted market returns at a daily frequency, but on a monthly basis it is non-significant to predict value-weighted market returns (Panel 2B) at the 5% level. More importantly, when employed together with CSV and CSK, it has a substantial decrease in its t-statistics but remains an important explanatory variable with exception of the results presented in Panel 2B to the value-weighted scheme.

Cross-sectional kurtosis (CSK) is also an important predictor of market returns, but it loses power when predicting value-weighted market returns in a daily frequency when taking into account CSV^{CW} and CSS with a t-statistics from 1,814 (until 2006) to 0,455 (until 2014), as shown in Panel 1B. Panel 2A reports that CSK is highly significant when considered alone, but becomes non-significant when other variables are included. In addition, Panel 2B shows that CSK has no power in any of the four analyzed regressions.

Also, to corroborate that the models with an equal-weight scheme seem more appropriate, it can be seen that the adjusted R^2 are higher for these regressions than their counterparts. Regarding Hypothesis 3 about the power of CSV in the presence of higher cross-sectional moments for the United States, it is verified that with exception of CSV^{CW} in Panel 1B, all the remaining estimations indicate that daily or monthly idiosyncratic variance matters in the presence of higher cross-sectional moments in the whole sample.

Table 17 - Predictability Regressions of Market Returns with Higher Cross-Sectional Moments - United States Sample

Panel 1A (Panel 1B) reports the daily predicitive regressions of the equal-weighted (value-weighted) market excess returns by lagged cross-sectional variance (CSV), skewness (CSS) and kurtosis (CSK). Panel 2A (2B) shows the monthly predicitive regressions of the equal-weighted (value-weighted) market excess returns by lagged cross-sectional variance (CSV), skewness (CSS) and kurtosis (CSK). Model 1 (2) uses only the CSS (CSK) as regressor. Model 3 uses all three variables jointly.The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) correct t-statistics in parentheses. The sample period is from July 1963 to December 2014.

	1963:7-2006:12					1963:7-2014:12				
	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$
Panel 1A: Forecasting r^{EW}										
1	0,001		0,003		5,8%	0,001		0,003		3,6%
t-stat	(5,390)		(20,478)			(5,443)		(20,485)		
2	-0,001			0,002	8,8%	-3,00E-04			0,002	5,2%
t-stat	(-4,041)			(18,792)		(-2,732)			(17,829)	
3	-0,001	0,426	0,001	0,002	9,6%	-0,001	0,285	0,001	0,001	5,7%
t-stat	(-5,820)	(5,037)	(3,612)	(12,004)		(-2,781)	(2,623)	(5,531)	(10,697)	
	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$
Panel 1B: Forecasting r^{CW}										
1	2,00E-04		0,002		1,1%	2,00E-04		0,002		0,8%
t-stat	(2,388)		(12,406)			(2,679)		(12,261)		
2	-2,00E-04			0,001	0,8%	-1,00E-04			5,00E-04	0,4%
t-stat	(-1,689)			(8,783)		(-0,506)			(6,891)	
3	-3,00E-04	1,151	0,001	2,00E-04	1,4%	-1,00E-04	0,945	0,002	1,00E-04	0,9%
t-stat	(-1,262)	(1,934)	(6,266)	(1,814)		(-0,675)	(1,895)	(7,708)	(0,455)	

Continued

Panel 2: Monthly	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$
Panel 2A: Forecasting r^{EW}										
1	0,007		0,003		4,9%	0,007		0,003		4,1%
t-stat	(3,118)		(5,685)			(3,079)		(5,677)		
2	-0,002			0,001	3,1%	-0,002			0,001	2,8%
t-stat	(-0,695)			(5,041)		(-0,490)			(4,809)	
3	-0,004	0,241	0,003	2,00E-04	5,6%	-0,005	0,239	0,003	3,00E-04	5,1%
t-stat	(-0,692)	(3,028)	(2,848)	(0,584)		(-0,922)	(2,611)	(2,309)	(0,864)	
	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$
Panel 2B: Forecasting r^{CW}										
1	0,005		0,001		0,3%	0,005		0,001		0,2%
t-stat	(2,390)		(1,620)			(2,386)		(1,671)		
2	0,004			1,00E-04	-0,1%	0,004			1,00E-04	-0,1%
t-stat	(1,416)			(0,656)		(1,460)			(0,654)	
3	0,012	-0,517	0,002	-3,00E-04	0,6%	0,014	-0,857	0,001	-3,00E-04	1,3%
t-stat	(3,032)	(-1,635)	(1,777)	(-1,370)		(4,252)	(-2,477)	(1,609)	(-1,149)	

Source: Author (2017)

Additionally, a more general model was tested, which includes to the higher cross-sectional moments a broader set of variables. Results are shown in Table 18.

Table 18 - Predictability Regressions of Market Returns by CSV Controlling for Market Variance, Investor Sentiment and Higher Cross-Sectional Moments – United States Sample

Panel A exhibits the monthly predicitive regressions of the equal-weighted market excess returns by the following lagged monthly variables: equal-weighted cross-sectional variance, cross-sectional-skewness, cross-sectional kurtosis, equal-weighted market variance, and investor sentiment. Panel B shows the results for the market returns, cross-sectional variance, and market variance value-weighted counterparts. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) correct t-statistics are shown in parentheses. The sample period is from July 1965 to December 2014.

Panel A: Equal-weighted measures							
	Constant	CSV^{EW}	CSS	CSK	MV^{EW}	$SENT$	$Adj.R^2$
Forecasting r^{EW}	-0,023	0,405	0,004	-2,00E-04	-2,258	-0,008	7,5%
t-stat	(-0,434)	(3,571)	(3,313)	(-0,680)	(-2,350)	(-3,767)	
Panel B: Cap-weighted measures							
	Constant	CSV^{CW}	CSS	CSK	MV^{CW}	$SENT$	$Adj.R^2$
Forecasting r^{CW}	0,015	-0,393	0,002	-0,001	-0,848	-0,003	1,6%
t-stat	(4,592)	(-1,052)	(2,024)	(-1,947)	(-1,879)	(-1,387)	

Source: Author (2017)

Table 18 shows that CSV^{EW}, CSS, and SENT are still relevant variables to predict one-month-ahead equal-weighted market returns while MV^{EW} is now significant in this specification. Additionally, CSV^{CW} loses significance while CSS and CSK have increased their relevance showing a statistical power indicating that value-weighted market variance and investor sentiment significantly reduce the contribution of aggregate idiosyncratic variance when compared to Panel 2B in Table 17 in the whole sample period.

The role of cross-sectional variance, skewness, and kurtosis were also investigated for the Brazilian case as described in Table 19.

Table 19 - Predictabily of Market Returns by with Higher Cross-Sectional Moments - Brazilian Sample

Panel 1A (Panel 1B) reports the daily predicitive regressions of the equal-weighted (value-weighted) market excess returns by lagged cross-sectional variance (CSV), skewness (CSS) and kurtosis (CSK). Panel 2A (2B) shows the monthly predicitive regressions of the equal-weighted (value-weighted) market excess returns by lagged cross-sectional variance (CSV), skewness (CSS) and kurtosis (CSK). Model 1 (2) uses only the CSS (CSK) as regressor. Model 3 uses all three variables jointly. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) correct t-statistics in parentheses. The sample period is from January 2000 to December 2016.

Panel 1: Daily

			2000-2016		
	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$
Panel 1A: Forecasting r^{EW}					
1	0,001		0,002		0,0%
t-stat	(2,623)		(1,699)		
2	4,70E-04			0,001	0,2%
t-stat	(1,993)			(2,944)	
3	0,001	-0,330	2,00E-04	0,001	0,6%
t-stat	(3,001)	(-1,996)	(0,205)	(2,837)	
	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$
Panel 1B: Forecasting r^{CW}					
1	1,20E-04		0,001		0,0%
t-stat	(0,512)		(0,495)		
2	9,50E-05			1,10E-04	0,0%
t-stat	(0,398)			(0,581)	
3	2,00E-04	-0,34	3,00E-04	1,00E-04	0,0%
t-stat	(0,381)	(-0,231)	(0,263)	(0,414)	

Panel 2: Monthly

	Constant	CSV^{EW}	CSS	CSK	$Adj.R^2$
Panel 2A: Forecasting r^{EW}					
1	0,002		0,012		26,2%
t-stat	(0,330)		(2,765)		
2	-0,008			0,002	7,6%
t-stat	(-1,340)			(4,119)	
3	-0,009	0,026	0,003	0,002	6,7%
t-stat	(-0,926)	(0,128)	(0,493)	(3,247)	

Continued

	Constant	CSV^{CW}	CSS	CSK	$Adj.R^2$
Panel 2B: Forecasting r^{CW}					
1	-0,002		0,008		1,1%
t-stat	(-0,405)		(1,780)		
2	-0,008			0,001	3,5%
t-stat	(-1,498)			(2,837)	
3	-0,008	-0,022	0,002	0,001	4,7%
t-stat	(-0,859)	(-0,029)	(0,376)	(2,178)	

Source: Author (2017)

The importance of cross-sectional skewness and kurtosis to the predictability of Brazilian market returns is interesting. In accordance with Table 19 and considering the equal-weighted market returns in a daily frequency, the results document that CSK is an important source of risk and robust to the inclusion of CSV^{EW} and CSS, which in turn supports the idea that higher risk predicts higher market returns. CSV^{EW} also becomes significant to predict one-day-ahead equal-weghted market returns in the presence of CSS and CSK. In opposition, these variables are not shown to be able to predict value-weighted market returns.

Testing the CSS and CSK in a monthly frequency gives somewhat different results in the sense that CSS, when considered by itself, is relevant in forecasting equal-weighted market returns at the 5% significance level and marginally significant in forecasting value-weighted market returns in contrary to what was observed in daily frequency. Perhaps even more interesting, CSK is relevant in each of the four models where it was added and retained its statistical power in the presence of CSV and CSS. These results indicate that the cross-section of kurtosis in the sample is more valuable than the skewness. Concerning Hypothesis 3 about the importance of CSV in the presence of higher cross-sectional moments, only in the case of Panel 1A, where CSV^{EW} is employed with daily data, the null hypothesis can be rejected, i.e, CSV^{EW} matters in the presence of cross-sectional skewness and kurtosis. Overall, aggregate idiosyncratic risk has no power to predict market in this setting with Brazilian data.

How sensitive results are to a more complete model specification which includes the market variance and investor sentiment was also tested. These results are shown in Table 20.

Table 20 - Predictability Regressions of Market Returns Controlling for CSV, Market Variance, Investor Sentiment and Higher Cross-Sectional Moments – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the following monthly lagged variables: equal-weighted cross-sectional variance, cross-sectional skewness, cross-sectional kurtosis, equal-weighted market variance, index of consumer confidence, gross domestic product, inflation and interest rate, respectively. Panel B shows the value-weighted versions of market returns, cross-sectional variance, and market variance in the predictive regressions. The main specification regression is defined as: $r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2003 to June 2016.

Panel A: Equal-weighted measures										
	Constant	CSV^{EW}	CSS	CSK	MV^{EW}	ICC	$IBC-Br$	$IPCA$	CDI	$Adj.R^2$
Forecasting r^{EW}	0,248	0,310	-0,004	0,002	-1,194	-1,00E-04	-0,002	-0,969	-3,256	12,8%
t-stat	(1,864)	(1,270)	(-0,661)	(2,435)	(-1,130)	(-0,181)	(-2,357)	(-0,579)	(-1,014)	
Panel B: Cap-weighted measures										
	Constant	CSV^{CW}	CSS	CSK	MV^{CW}	ICC	$IBC-Br$	$IPCA$	CDI	$Adj.R^2$
Forecasting r^{CW}	0,144	2,441	-0,002	0,001	-2,320	-1,00E-04	-0,001	-1,477	-1,842	5,4%
t-stat	(1,523)	(1,102)	(-0,443)	(1,472)	(-1,244)	(-0,270)	(-2,058)	(-1,281)	(-0,764)	

Source: Author (2017)

Table 20 shows the results when using aggregate idiosyncratic variance to predict one-month ahead market returns controlling for other predictors that may influence the relation. It is observed that when forecasting equal-weighted market returns, CSK is still a significant variable implying that higher cross-sectional kurtosis gives a higher equal-weighted market return for the next period. However, CSK becomes non-significant when these additional variables were included regarding the value-weighted measures. Independently from the weighting scheme, the IBC-Br is shown to play an important role as it continues to be significant. More importantly, these results support the idea that aggregate idiosyncratic variance does not forecast monthly market returns.

4.1.7 Predictability of Risk Factors and Portfolios based on Realized Idiosyncratic Volatility using Aggregate Idiosyncratic Variance

Another investigation pursued in this work was trying to understand if CSV can be seen as a state variable able to predict omitted risk factors from the CAPM model. This subsection reviews how CSV performs in predicting risk factors used in Carhart's (1997) model and the recently proposed five-factor model of Fama and French (2015). Table 21 shows the results for the United States stock market.

Table 21 - Forecasts of Risk Factors by CSV and Market Variance – United States Sample

This table shows the results of a one-month-ahead predictive regression of the risk factors in Carhart's (1997) model on the monthly lagged equal-weighted cross-sectional variance and lagged equal-weighted market variance (Panel 1A) and on the monthly value-weighted cross-sectional variance and lagged value-weighted market variance (Panel 1B). Panel 2 exhibits the predictive regressions based on the Fama and French (2015) five factor model. Panel 2A shows the estimation based on equal-weighted scheme and Panel 2B on value-weighted scheme. The estimated regression is: $r_{t+1} = a + \beta v_t + \gamma m v_t + \varepsilon_{t+1}$, where r_{t+1} represents the risk factor under analysis in period t+1, a is the constant of the model, β is the slope coefficient related to the aggregate idiosyncratic risk, v_t represents either the equal or value-weighted idiosyncratic variance in period t, γ is the slope coefficient related to the market variance, mv_t is either the equal or value-weighted market variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from July 1963 to December 2014.

Panel 1: Carhart (1997) risk factors

Panel 1A: Equal-weighted measures

Equation	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
SMB	1,00E-04	0,068	-0,132	0,0%
t-stat	(0,025)	(1,286)	(-0,404)	
HML	0,004	0,033	-0,941	0,7%
t-stat	(1,723)	(0,408)	(-1,752)	
Momentum	0,017	-0,247	-0,946	3,9%
t-stat	(3,250)	(-1,458)	(-0,952)	

Panel 1B: Cap-weighted measures

Equation	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
SMB	0,002	0,059	0,069	-0,3%
t-stat	(0,730)	(0,185)	(0,189)	
HML	0,005	0,823	-1,466	2,4%
t-stat	(0,161)	(1,493)	(-2,099)	
Momentum	0,011	-0,306	-0,722	0,8%
t-stat	(4,665)	(-0,915)	(-0,998)	

Panel 2: Fama and French (2015) risk factors

Panel 2A: Equal-weighted measures

Equation	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
SMB	2,00E-04	0,805	-0,279	0,1%
t-stat	(0,086)	(1,356)	(-0,752)	
HML	0,004	0,033	-0,941	0,7%
t-stat	(1,723)	(0,408)	(-1,752)	
CMA	0,002	0,022	0,157	-0,1%
t-stat	(1,061)	(0,374)	(0,396)	
RMW	0,003	0,007	-0,021	-0,3%
t-stat	(1,964)	(0,141)	(-0,085)	

Continued

Panel 2B: Cap-weighted measures				
Equation	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
SMB	0,001	0,252	-0,218	-0,2%
t-stat	(0,551)	(0,918)	(-0,615)	
HML	0,005	0,823	-1,466	2,4%
t-stat	(0,161)	(1,493)	(-2,099)	
CMA	-0,001	0,530	-0,379	0,9%
t-stat	(-0,296)	(0,990)	(-0,665)	
RMW	-0,001	0,728	-0,612	2,4%
t-stat	(-0,745)	(2,871)	(-2,231)	

Source: Author (2017)

Table 21 shows the predictive power of aggregate idiosyncratic variance and market variance for risk factors in the Carhart (1997) and Fama and French (2015) five-factor model. The analysis here point out that aggregate idiosyncratic variance and market variance cannot satisfactorily predict the size, value and momentum premium in Carhart's (1997) model.

Although not significant, the signs for the value and momentum premium are positive and negative, respectively, in accordance with Angelidis, Sakkas, and Tessaromatis (2015), which use an international sample of the G7 countries and a cross-sectional dispersion measure, though constructed differently from Garcia, Mantilla-García, and Martellini (2014) (see Angelidis, Sakkas, and Tessaromatis, 2015). Additionally, the authors find that return dispersion is not highly correlated with a measure of idiosyncratic volatility based on the residuals of the CAPM model, then it has a different behavior of what is shown to the United States case and may shed some light regarding why these findings differ in terms of statistical significance.

Also, for quarterly data and using Goyal and Santa-Clara's (2003) measure, Guo and Savickas (2006) find value and momentum premium to be negatively related to market variance, while idiosyncratic volatility was only (positively) significant to forecast the HML factor. Moreover, Angelidis, and Tessaromatis (2008) show using a sample for the United Kingdom that the value-weighted idiosyncratic volatility is not significant to forecast size premium although the equal-weighted idiosyncrantic risk (market variance) is positive (negative) at the 5% level. The authors also show that neither idiosyncratic volatility nor the market variance have the ability to forecast value premium.

The new evidence on the five-factors models reflect that only market variance negatively predicts the value premim and both together predict the profitability factor (RMW) using a value-weighted scheme; thus, higher idiosyncratic variance generates a higher premimum for the stocks with higher profitability (robust) in relation to those with lower profitability (weak).

Table 22 now focuses on aggregate idiosyncratic variance's and market variance's ability to predict the risk factors for the Brazilian stocks.

Table 22 - Forecasts of Risk Factors by CSV and Market Variance – Brazilian Sample

The table shows the results of a one-month-ahead predictive regression of the risk factors in Carhart's (1997) model on the monthly lagged equal-weighted cross-sectional variance and lagged equal-weighted market variance (Panel 1A) and on the monthly lagged value-weighted cross-sectional variance and lagged value-weighted market variance (Panel 1B). Panel 2 exhibits the predictive regressions based on the Fama and French (2015) five factor model. Panel 2A shows the estimations based on equal-weighted scheme and Panel 2B on value-weighted scheme. The estimated regression is: $r_{t+1} = a + \beta v_t + \gamma m v_t + \varepsilon_{t+1}$, where r_{t+1} represents the risk factor under analysis in period t+1, a is the constant of the model, β is the slope coefficient related to the aggregate idiosyncratic risk, v_t represents either the equal or value-weighted idiosyncratic variance in period t, γ is the slope coefficient related to the market variance, mv_t is either the equal or value-weighted market variance in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

Panel 1: Carhart (1997) risk factors

Panel 1A: Equal-weighted measures

Equation	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
SMB	0,002	-0,041	-0,922	0,8%
t-stat	(0,446)	(-0,332)	(-1,768)	
HML	-0,007	-0,002	1,114	0,3%
t-stat	(-1,060)	(-0,011)	(1,446)	
Momentum	0,014	0,022	-0,634	-0,6%
t-stat	(2,082)	(0,135)	(-0,845)	

Panel 1B: Cap-weighted measures

Equation	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
SMB	-0,006	1,201	-1,539	1,2%
t-stat	(-0,762)	(1,061)	(-1,753)	
HML	-0,015	1,472	-0,220	1,0%
t-stat	(-1,835)	(1,178)	(-0,213)	
Momentum	0,026	-2,012	0,734	1,4%
t-stat	(3,346)	(-1,665)	(0,735)	

Continued

Panel 2: Fama and French (2015) risk factors				
Panel 2A: Equal-weighted measures				
Equation	Constant	CSV^{EW}	MV^{EW}	$Adj.R^2$
SMB	0,006	-0,019	-0,578	-0,4%
t-stat	(1,063)	(-0,144)	(-0,951)	
HML	-0,007	-0,002	1,114	0,3%
t-stat	(-1,060)	(-0,011)	(1,446)	
CMA	-0,01	0,192	0,256	0,4%
t-stat	(-1,739)	(1,275)	(0,377)	
RMW	0,007	0,030	-0,342	-0,9%
t-stat	(1,057)	(0,142)	(-0,731)	
Panel 2B: Cap-weighted measures				
Equation	Constant	CSV^{CW}	MV^{CW}	$Adj.R^2$
SMB	-0,003	1,526	-1,536	0,8%
t-stat	(-0,524)	(1,551)	(-1,890)	
HML	-0,015	1,472	-0,220	1,0%
t-stat	(-1,835)	(1,178)	(-0,213)	
CMA	-0,014	1,838	-0,993	0,6%
t-stat	(-1,861)	(1,580)	(-0,872)	
RMW	0,013	-0,873	0,303	-0,5%
t-stat	(1,845)	(-0,775)	(0,325)	

Source: Author (2017)

Table 22 analyzes the predictability capacity of idiosyncratic variance and market variance to forecast market returns. Here becomes evident that regardless the weighting scheme applied and considering the conventional 5% of significance level none of the models estimated exhibited that idiosyncratic variance helps to predict the risk factors.

With 10% of significance level and the Carhart (1997) model only MV^{EW} matters to predict the SMB factor in Panel 1A while CSV^{EW} is non-significant. The same happens in Panel 1B with the value-weighted scheme in addition to the CSV^{CW} to predict the momentum factor, but then it is the MV^{CW} which has no statistical power. In relation to the Fama and French (2015) five-factor model, Panel 2A reports that again none of the variables are relevant to predict the risk factors whereas Panel 2B shows only marginal significance of the value-weighted market variance to predict the size premium, however idiosyncratic variance remains non-significant.

In general, concerning Hypothesis 4 about the CSV ability to predict risk factors, for the United States sample the analyses indicate that only CSV^{CW} is relevant to predict the profitability factor (RMW) and controlling for MV^{CW} in the model' specification, i.e, very limited evidence is reported here. For the Brazilian case, CSV^{EW} or CSV^{CW} does not play a relevant role in predicting these risk factors and taking into account their corresponding market variance.

Additionally, it may be of importance to evaluate whether idiosyncratic variance helps predict the return of portfolios based on the lagged realized idiosyncratic volatility (IVOL). The results are shown in Table 23.

Table 23 - Forecasts of Realized Idiosyncratic Volatility Portfolios by CSV and Market Variance – Brazilian Sample

This table shows the results of a one-month-ahead predictive regression of the value-weighted portfolio excess returns based on realized idiosyncratic volatility as in Ang *et al.* (2006) on the monthly lagged value-weighted cross-sectional variance and lagged value-weighted market variance. High represents the portfolios containing stocks at the highest idiosyncratic volatility and Low with the lowest idiosyncratic volatility. The estimated regression is: $r_{pt+1} = a + \beta v_t + \gamma m v_t + \varepsilon_{t+1}$, where r_{pt+1} represents the excess return on the value-weighted portfolio returns formed according to the stock's realized idiosyncratic volatility in period t+1; a is the constant of the model, β is the slope coefficient related to the aggregate idiosyncratic risk, v_t represents the value-weighted idiosyncratic variance in period *t*, γ is the slope coefficient related to the market variance, mv_t is the value-weighted market variance in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

Equation	Constant	CSV^{CW}	MV^{CW}	$Adj. R^2$	Mean	Standard deviation
High	-0,010	0,677	-0,991	-0,5%	0,2%	9,4%
t-stat	(-0,703)	(0,293)	(-0,522)			
2	-0,01	3,099	-2,583	0,3%	1,6%	8,3%
t-stat	(-0,621)	(1,094)	(-1,050)			
3	-0,01	1,486	-1,161	-0,6%	0,9%	7,0%
t-stat	(-0,679)	(0,607)	(-0,602)			
4	-0,004	1,373	-1,082	-0,6%	1,4%	6,3%
t-stat	(-0,303)	(0,689)	(-0,628)			
Low	-4,00E-04	1,184	-1,822	0,4%	1,2%	6,6%
t-stat	(-0,030)	(0,488)	(-0,789)			

Source: Author (2017)

The realized idiosyncratic volatility was used by Ang *et al.* (2006, 2009) to verify the relation between idiosyncratic risk and expected returns and they found a statistical negative association between them. Guo and Savickas (2006) also verify if the Goyal and Santa-Clara (2003) measure, which is mostly idiosyncratic, using quarterly data, is able to forecast these portfolio returns together with market variance. Therefore, this was evaluated to the cross-sectional variance in the Brazilian stock market.

Table 23 shows the ability of cross-sectional variance and market variance in predicting portfolio returns based on the stock's realized idiosyncratic volatility. The results for the Brazilian case show that it is not possible to predict portfolios returns based on realized idiosyncratic variance with neither $\mathrm{CSV}^{\mathrm{CW}}$ nor $\mathrm{MV}^{\mathrm{CW}}$, which is different from what was found by Guo and Savickas (2006) with American data.

However, there is another interesting point in Table 23 concerning the behavior of the quintile portfolios used in the estimation: as it is widely seen in the literature in the United States stock market and pointed out as controversial, a portfolio with a lower level of risk has a higher mean return than that of a higher risk portfolio, which is in agreement with the results provided by Ang *et al.* (2006) and gives rise to the idiosyncratic volatility puzzle. The same pattern emerged from the sample used in the current work with stocks traded on Brazilian stock market.

4.1.8 Aggregate Idiosyncratic Variance by Levels of Credit Ratings and Size

The question of whether idiosyncratic variance proxied by cross-sectional variance constructed only from stocks with available credit ratings at the firm level can be useful to forecast market returns is addressed. In order to have a better understanding of the idiosyncratic risk of rated firms, the sample of stocks is split into three parts: stocks which belong to firms rated as non-investment grade (NIG), investment grade (IG) and all rated firms for the United States sample. Table 24 shows the results.

Table 24 -Daily Predictability of Market Returns using CSV by Levels of Credit Ratings – United States Sample

Panel A shows the results of one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance based on non-investment grade (NIG) in Panel A, investment grade (IG) in Panel C and all rated firms (RAT) in Panel E for four sample periods. Panels B, D and F present the value-weighted daily market excess returns on the daily lagged value-weighted counterparts. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on NIG, IG, or RAT firms in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 1987 to December 2014.

	1987:1 - 1999:12	1987:1 - 2001:12	1987:1 - 2006:12	1987:1 - 2014:12
	CSV_NIG^{EW}	CSV_NIG^{EW}	CSV_NIG^{EW}	CSV_NIG^{EW}
Panel A: Forecasting r^{EW}				
Constant	2,12E-04	6,51E-05	0,001	0,001
NW t - statistics	(0,670)	(0,203)	(5,639)	(4,208)
Coefficient	0,401	0,417	0,018	0,063
NW t - statistics	(1,825)	(2,087)	(0,627)	(0,999)
$Adj.R^2$	0,4%	0,5%	0,0%	0,0%
	CSV_NIG^{CW}	CSV_NIG^{CW}	CSV_NIG^{CW}	CSV_NIG^{CW}
Panel B: Forecasting r^{CW}				
Constant	-4,96E-04	-2,80E-05	5,38E-05	-1,64E-05
NW t - statistics	(-1,183)	(-0,081)	(0,196)	(-0,067)
Coefficient	1,223	0,370	0,292	0,428
NW t - statistics	(2,450)	(1,103)	(0,985)	(1,425)
$Adj.R^2$	0,5%	0,1%	0,0%	0,1%
	CSV_IG^{EW}	CSV_IG^{EW}	CSV_IG^{EW}	CSV_IG^{EW}
Panel C: Forecasting r^{EW}				
Constant	0,001	0,001	0,001	4,84E-04
NW t - statistics	(1,572)	(2,145)	(2,743)	(2,351)
Coefficient	0,712	0,430	0,357	0,702
NW t - statistics	(0,509)	(0,454)	(0,442)	(1,159)
$Adj.R^2$	0,1%	0,0%	0,0%	0,1%

Continued

Panel D: Forecasting r^{CW}	CSV_IG^{CW}	CSV_IG^{CW}	CSV_IG^{CW}	CSV_IG^{CW}
Constant	-0,001	-2,57E-04	-1,83E-04	4,43E-05
NW t - statistics	(-1,814)	(-0,583)	(-0,503)	(0,174)
Coefficient	4,376	1,821	1,691	0,965
NW t - statistics	(3,655)	(1,473)	(1,493)	(1,186)
$Adj.R^2$	1,3%	0,4%	0,3%	0,1%
Panel E: Forecasting r^{EW}	CSV_RAT^{EW}	CSV_RAT^{EW}	CSV_RAT^{EW}	CSV_RAT^{EW}
Constant	1,61E-04	1,27E-04	0,001	0,001
NW t - statistics	(0,376)	(0,353)	(5,669)	(3,776)
Coefficient	0,799	0,691	0,036	0,140
NW t - statistics	(1,484)	(1,687)	(0,601)	(1,035)
$Adj.R^2$	0,4%	0,4%	0,0%	0,1%
Panel F: Forecasting r^{CW}	CSV_RAT^{CW}	CSV_RAT^{CW}	CSV_RAT^{CW}	CSV_RAT^{CW}
Constant	-0,001	-2,34E-04	-1,66E-04	3,41E-06
NW t - statistics	(-1,980)	(-0,524)	(-0,446)	(0,013)
Coefficient	4,031	1,495	1,387	0,945
NW t - statistics	(3,802)	(1,384)	(1,395)	(1,258)
$Adj.R^2$	1,3%	0,3%	0,2%	0,1%

Source: Author (2017)

Overall, the results in Table 24 show that the aggregate idiosyncratic variance using a daily frequency and based in accordance with the credit ratings do not play a role in forecasting market returns across the sample period analyzed. However, the importance of value-weighted idiosyncratic variance for the period January 1987 to December 1999 is shown to be robust across the credit ratings categories but not along the years. The results for the monthly frequency are shown in Table 25 for firms rated as non-investment grade (*NIG*).

Table 25 - Monthly Predictability of Market Returns using CSV based on Non-Investment Grade Firms – United States Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance based on firms rated as non-investment grade (NIG) in Panel A for four sample periods. Panel B shows the value-weighted monthly market excess returns on the monthly lagged value-weighted counterpart. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient; v_t represents either the equal or value-weighted idiosyncratic variance based on firms rated as non-investment grade (NIG) in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 1987 to December 2014.

	1987:1 - 1999:12	1987:1 - 2001:12	1987:1 - 2006:12	1987:1 - 2014:12
	CSV_NIG^{EW}	CSV_NIG^{EW}	CSV_NIG^{EW}	CSV_NIG^{EW}
Panel A: Forecasting r^{EW}				
Constant	-0,021	-0,020	0,003	0,003
NW t - statistics	(-2,570)	(-2,296)	(0,543)	(0,560)
Coefficient	0,762	0,673	0,127	0,141
NW t - statistics	(3,818)	(3,242)	(0,913)	(0,820)
$Adj.R^2$	6,5%	4,7%	0,0%	0,2%
	CSV_NIG^{CW}	CSV_NIG^{CW}	CSV_NIG^{CW}	CSV_NIG^{CW}
Panel B: Forecasting r^{CW}				
Constant	0,001	0,012	0,012	0,014
NW t - statistics	(0,163)	(2,001)	(2,841)	(4,212)
Coefficient	0,478	-0,271	-0,354	-0,498
NW t - statistics	(0,918)	(-0,910)	(-1,434)	(-2,180)
$Adj.R^2$	0,1%	0,0%	0,6%	1,6%

Source: Author (2017)

Table 25 shows that CSV_NIG^{EW} and CSV_NIG^{CW} were important only in three cases: in the first two periods of the sample from January 1987 to December 2001 to forecast equal-weighted market returns; and in the third during the whole period, from January 1987 to December 2014, respectively, to forecast value-weighted market returns. Interestingly, the first two have a significantly positive sign while the last is significantly negative indicating a different behavior between the variables and along the years. Table 26 shows the results for firms rated as investment grade (IG).

Table 26 - Monthly Predictability of Market Returns using CSV based on Investment Grade Firms – United States Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance based on firms rated as investment grade (IG) in Panel A for four sample periods. Panel B shows the value-weighted monthly market excess returns on the monthly lagged value-weighted counterpart. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on firms rated as investment grade (IG) in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 1897 to December 2014.

	1987:1 - 1999:12	1987:1 - 2001:12	1987:1 - 2006:12	1987:1 - 2014:12
	CSV_IG^{EW}	CSV_IG^{EW}	CSV_IG^{EW}	CSV_IG^{EW}
Panel A: Forecasting r^{EW}				
Constant	-0,007	-0,001	0,004	0,011
NW t - statistics	(-0,812)	(-0,083)	(0,621)	(1,481)
Coefficient	1,907	0,846	0,540	-0,380
NW t - statistics	(1,470)	(0,849)	(0,662)	(-0,355)
$Adj.R^2$	1,5%	0,0%	-0,2%	-0,1%
	CSV_IG^{CW}	CSV_IG^{CW}	CSV_IG^{CW}	CSV_IG^{CW}
Panel B: Forecasting r^{CW}				
Constant	0,009	0,015	0,013	0,016
NW t - statistics	(1,289)	(3,192)	(3,401)	(4,360)
Coefficient	0,009	-1,268	-1,153	-1,601
NW t - statistics	(0,006)	(-1,956)	(-1,939)	(-2,383)
$Adj.R^2$	-0,7%	1,5%	1,2%	3,7%

Source: Author (2017)

Using investment grade firms, the results were different from the previous exercise. In this case, only CSV_IG^{CW} is shown to play a role in predicting value-weighted market returns. In two out of four periods analyzed (January 1987 until December 2006), it was marginally significant, whereas it was statiscally significant at a 5% level for one out of four, which is the whole sample. Additionally, CSV_IG^{CW} still has a negative sign in all three cases as in the *NIG* firms. Finally, the analysis now turns to all rated firms in Table 27.

Table 27 - Monthly Predictability of Market Returns using CSV based on All Rated Firms – United States Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance based on all rated firms (RAT) for four sample periods. Panel B shows the value-weighted monthly market excess returns on the monthly lagged value-weighted counterpart. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on all rated firms (RAT) in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 1987 to December 2014.

	1987:1 - 1999:12	1987:1 - 2001:12	1987:1 - 2006:12	1987:1 - 2014:12
	CSV_RAT^{EW}	CSV_RAT^{EW}	CSV_RAT^{EW}	CSV_RAT^{EW}
Panel A: Forecasting r^{EW}				
Constant	-0,020	-0,015	0,004	0,005
NW t - statistics	(-2,438)	(-1,672)	(0,628)	(0,760)
Coefficient	1,379	0,962	0,209	0,155
NW t - statistics	(3,196)	(2,362)	(0,882)	(0,439)
$Adj.R^2$	5,5%	3,0%	-0,1%	-0,1%
	CSV_RAT^{CW}	CSV_RAT^{CW}	CSV_RAT^{CW}	CSV_RAT^{CW}
Panel B: Forecasting r^{CW}				
Constant	0,009	0,015	0,013	0,016
NW t - statistics	(1,189)	(3,245)	(3,578)	(4,577)
Coefficient	0,068	-1,077	-1,031	-1,409
NW t - statistics	(0,052)	(-2,029)	(-2,124)	(-2,537)
$Adj.R^2$	-0,7%	0,0%	0,0%	3,60%

Source: Author (2017)

Table 27 exhibits a similar pattern to what was shown in the previous two tables. Specifically, CSV_RAT^{EW} is positively and statistically significant during the first two periods analyzed, i.e. until December 2001, despite the negative and significant sign of CSV_RAT^{CW} with exception of the first period (until 1999). In light of these findings, because equal-weighted measures are driven by small firms, it seems that it was the idiosyncratic variance of small firms and firms rated as non-investment grade rating that were useful to forecast equal-weighted market returns, but only until December 2001.

In opposition, because a market capitalization scheme gives more weight to bigger firms, it seems that it is the aggregate risk of big firms within an investment-grade category that is more important to negatively forecast value-weighted market returns in a more recent period when taking into account all rated firms. Therefore, in relation to Hypothesis 5 about the importance of a measure of idiosyncratic variance based on overall firm's credit rating and taking into account the whole period analyzed, idiosyncratic variance based on NIG, IG, and all rated firms do not forecast daily equal- or value-weighted market returns, but monthly CSV^{CW} according to each of these three categories matter, hence, the null hypothesis can only be rejected when using a value-weighted idiosyncratic variance measure according to NIG, IG, and all rated firms on a monthly basis.

Furthermore, trying to get a better sense of whether downgrade periods affect the relation between idiosyncratic variance and market returns, Avramov *et al.*'s (2013) procedure was followed: every month, t the stocks belonging to downgraded firms are excluded from the analysis as well as their return observations in the six months before and after the month t. This process was repeated every month in the sample. The authors found that removing downgraded periods makes the return on various anomalies-based strategies lose significance. This means that it was this distress period that was generating the abnormal returns. Therefore, this research intended to verify if downgraded periods have any influence on market return – aggregate idiosyncratic variance association. Table 28 summarizes the results.

Restricting the analysis to daily data only CSV_RAT^{EW} is significant and positive as opposed to the results shown in Table 24 (Panel E), which demonstrate that it had no forecasting power. It means that is the aggregate idiosyncratic risk of all rated firms using an equal-weighted scheme in stable or improving periods that is relevant to forecast equal-weighted market returns. Perhaps more interestingly, excluding downgraded periods has no effect in the power and sign of CSV_RAT^{EW} and CSV_RAT^{CW} in monthly data, as it can be seen that the former remains insignificant in the complete timespan and the latter is significant and negative. Overall, only daily CSV_RAT^{EW} was affected when the downgraded periods were removed.

Table 28 - Monthly Predictability of Market Returns using CSV based on All Rated Firms and Excluding Downgraded Periods – United States Sample

Each month *t* all downgraded firms were excluded from the sample (and also for the previous and next six months). This table shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance based on all rated firms (RAT) in Panel 1A. Panel 1B shows the value-weighted scheme. Panel 2 shows the corresponding predictive regressions using monthly data. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on all rated firms (RAT) in period *t*, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 1987 to December 2014.

	Constant	CSV_RAT^{EW}	CSV_RAT^{CW}	$Adj.R^2$
Panel 1: Daily				
Panel 1A: Forecasting r^{EW}	1,00E-04	0,896		0,3%
t-stat	(0,327)	(2,345)		
Panel 1B: Forecasting	-2,00E-05		1,116	0,1%
t-stat	(-0,074)		(1,232)	
Panel 2: Monthly				
	Constant	CSV_RAT^{EW}	CSV_RAT^{CW}	$Adj.R^2$
Panel 2A: Forecasting r^{EW}	0,004	0,250		-0,1%
t-stat	(0,514)	(0,439)		
Panel 2B: Forecasting r^{CW}	0,016		-1,494	2,8%
t-stat	(4,945)		(-2,772)	

Source: Author (2017)

The next empirical evidences shown in this subsection refer to the use of CSV by levels of credit ratings and applied to the stock's market capitalization due to the severe limitation of credit ratings information using Brazilian data. Also, as previously mentioned, the sample of firms with a non-investment grade rating level in the sample differs from the investment grade only from December 2004. Therefore, results concerning the CSV by rating levels are summarized first in Table 29 below, followed by those of size.

Table 29 - Daily Predictabily of Market Returns using CSV by levels of Credit Ratings – Brazilian Sample

Panel A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance based on all firms rated as non-investment grade (NIG) and all rated (RAT) firms for three sample periods. Panel B shows the value-weighted daily market excess returns on the daily lagged value-weighted counterparts. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on NIG or RAT firms in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000 - 2016	2000-2007	2008-2016	2000 - 2016	2000-2007	2008-2016
	$CSV - NIG^{EW}$	$CSV - NIG^{EW}$	$CSV - NIG^{EW}$	$CSV - RAT^{EW}$	$CSV - RAT^{EW}$	$CSV - RAT^{EW}$
Panel A. Forecasting r^{EW}						
Constant	0,001	0,001	0,000	0,001	0,001	1,00E-04
NW t - statistics	(1,834)	(2,079)	(0,265)	(1,940)	(2,088)	(0,407)
Coefficient	0,202	0,492	0,146	0,190	0,492	0,107
NW t - statistics	(0,937)	(1,508)	(0,615)	(0,830)	(1,738)	(0,426)
$Adj.R^2$	0,0%	0,1%	0,0%	0,0%	0,1%	0,0%
	$CSV - NIG^{CW}$	$CSV - NIG^{CW}$	$CSV - NIG^{CW}$	$CSV - RAT^{CW}$	$CSV - RAT^{CW}$	$CSV - RAT^{CW}$
Panel B. Forecasting r^{CW}						
Constant	-1,00E-04	2,00E-04	-0,001	-0,001	3,00E-04	-0,001
NW t - statistics	(-0,341)	(0,515)	(-0,910)	(-1,603)	(0,739)	(-2,628)
Coefficient	0,588	0,511	0,736	1,927	0,282	3,371
NW t - statistics	(0,648)	(0,607)	(0,613)	(2,171)	(0,326)	(2,472)
$Adj.R^2$	0,0%	0,0%	0,0%	0,2%	-0,1%	0,5%

Source: Author (2017)

A very interesting point that can be extracted from the daily predictability regressions segmented by ratings is that $\mathrm{CSV\_RAT}^{\mathrm{CW}}$ is different from zero as opposed to the $\mathrm{CSV\_NIG}^{\mathrm{CW}}$, $\mathrm{CSV\_NIG}^{\mathrm{EW}}$, $\mathrm{CSV\_RAT}^{\mathrm{EW}}$ in the whole sample period. This evidence suggests that it is the value-weighted aggregate idiosyncratic variance based on all Brazilian rated firms in the sample which matters to forecast daily (value-weighted) market returns. Table 30 describes the findings with monthly frequency[12].

[12] The correlation between value-weighted aggregate idiosyncratic variance measured using the CSV for firms rated as NIG is shown to be greater than 93%, 90%, and 89% with alternative measures based on the multifactor models (FF-3, FFC, and FF-5), Bali, Cakici and Levy (2008), and Goyal and Santa-Clara (2003), respectively; for the equal-weighted idiosyncratic risk, the correlation values are greater than 96%,77%, and 90%, respeticvely. Considering the sample of all rated firms, RAT, the correlation for the value-weighted measures between the CSV and those based on the based on the multifactor models (FF-3, FFC, and FF-5), Bali, Cakici and Levy (2008), and Goyal and Santa-Clara (2003) are greater than 89%, 89%, and 78%, respectively; for the equal-weighted version are greater than 97%, 77%, and 89%. These values show that, in general, CSV is highly correlated with other well-used approaches.

Table 30 - Monthly Predictabily of Market Returns using CSV by levels of Credit Ratings – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance based on all firms rated as non-investment grade (NIG) and all rated firms (RAT) for three sample periods. Panel B uses value-weighted monthly market excess returns on the monthly lagged value-weighted counterparts. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on NIG or RAT firms in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000-2016	2000-2007	2008-2016	2000-2016	2000-2007	2008-2016
	$CSV-NIG^{EW}$	$CSV-NIG^{EW}$	$CSV-NIG^{EW}$	$CSV-RAT^{EW}$	$CSV-RAT^{EW}$	$CSV-RAT^{EW}$
Panel A: Forecasting r^{EW}						
Constant	0,006	0,025	-0,008	0,003	0,026	-0,011
NW t - statistics	(0,684)	(1,328)	(-0,839)	(0,378)	(1,355)	(-1,093)
Coefficient	0,064	-0,674	0,362	0,253	-0,813	0,541
NW t - statistics	(0,160)	(-0,491)	(0,942)	(0,606)	(-0,551)	(1,540)
$Adj.R^2$	-0,5%	-0,6%	-0,3%	-0,3%	-0,5%	0,5%
	$CSV-NIG^{CW}$	$CSV-NIG^{CW}$	$CSV-NIG^{CW}$	$CSV-RAT^{CW}$	$CSV-RAT^{CW}$	$CSV-RAT^{CW}$
Panel B: Forecasting r^{CW}						
Constant	0,009	0,036	-0,005	0,009	0,037	-0,005
NW t - statistics	(1,195)	(2,107)	(-0,583)	(1,044)	(2,184)	(-0,518)
Coefficient	-0,809	-3,442	-0,018	-1,044	-3,674	-0,063
NW t - statistics	(-1,270)	(-1,809)	(-0,023)	(-1,077)	(-1,891)	(-0,051)
$Adj.R^2$	0,8%	2,4%	-1,0%	0,0%	2,7%	-1,0%

Source: Author (2017)

Again, in general the results are in conformity with the general case when aggregate idiosyncratic variance was based on all stocks independent of whether the firm which the stock is linked to has an assigned credit rating or not. In other words, the findings show that none of the estimated models for both equal- and value-weighted schemes split into credit ratings levels reveal that these aggregate idiosyncratic variance measures are relevant to forecast market returns with the adjusted R^2 mainly negative in the performed regressions reflecting that these variables do not adjust well to the purpose of predicting market returns. In summary, Hypothesis 5 which refers to the relevance of idiosyncratic variance measures based on the firm's overall credit rating using the whole time period show that for the Brazilian case the null hypothesis cannot be rejected with exception of the relevance of daily $CSV - RAT^{CW}$ to forecast daily value-weighted market returns none of the other models suggest these measures matter. Disentangling the CSV by size is shown in the next table.

Table 31 - Daily Predictability Regressions of Market Returns using CSV by Size – Brazilian Sample

Panel A shows the results of a one-day-ahead predictive regression of the equal-weighted daily market excess returns on the daily lagged equal-weighted cross-sectional variance based on small (SMALL) and big (BIG) according to the stock's market capitalization for three sample periods. Panel B shows the value-weighted daily market excess returns on the daily lagged value-weighted counterparts. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on SMALL and BIG stocks in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000 - 2016	2000-2007	2008-2016	2000 - 2016	2000-2007	2008-2016
	$CSV - SMALL^{EW}$	$CSV - SMALL^{EW}$	$CSV - SMALL^{EW}$	$CSV - BIG^{EW}$	$CSV - BIG^{EW}$	$CSV - BIG^{EW}$
Panel A: Forecasting r^{EW}						
Constant	0,001	0,002	0,001	0,001	1,42E-05	0,001
NW t - statistics	(3,357)	(3,734)	(1,280)	(3,028)	(0,196)	(1,828)
Coefficient	-0,149	-0,133	-0,171	-0,541	2,018	-0,695
NW t - statistics	(-1,805)	(-1,740)	(-1,058)	(-1,236)	(1,530)	(-2,101)
$Adj.R^2$	0,2%	0,2%	0,2%	0,2%	0,3%	0,8%
	$CSV - SMALL^{CW}$	$CSV - SMALL^{CW}$	$CSV - SMALL^{CW}$	$CSV - BIG^{CW}$	$CSV - BIG^{CW}$	$CSV - BIG^{CW}$
Panel B: Forecasting r^{CW}						
Constant	0,001	4,00E-04	3,00E-04	-3,05E-06	4,50E-04	-4,32E-04
NW t - statistics	(2,288)	(0,757)	(0,821)	(-0,006)	(1,053)	(-0,479)
Coefficient	-0,412	0,056	-0,421	0,323	-0,054	0,700
NW t - statistics	(-1,112)	(0,131)	(-1,561)	(0,259)	(-0,081)	(0,299)
$Adj.R^2$	0,9%	-0,1%	1,7%	0,0%	-0,1%	0,0%

Source: Author (2017)

Interestingly, for the group of stocks with low market capitalization (SMALL) as a whole, only CSV_SMALL^{EW} is marginally significant (negative) for the complete period and the first part of the sample (until December 2007). For the group of stocks whith high market capitalization (BIG), it can be seen that CSV_BIG^{EW} is negative and significant for the second part of the sample (from January 2008), but undistinguishable from zero if only the complete period is taken into account. However, for the value-weighted scheme, no estimated model seems relevant to predict market returns. Also, the adjusted R^2 are very low reflecting that not many of the observed values can be predicted in all estimated models raging from 0,2% to 0,8% with an equal-weighted scheme and -0,1% to 1,7% with the value-weighted idiosyncratic variance measures. Next, monthly results are shown in Table 32[13].

[13] The correlation between value-weighted aggregate idiosyncratic variance measured using the CSV for stocks classified as big is shown to be greater than 92%, 98%, and 90% with alternative measures based on the multifactor models (FF-3, FFC, and FF-5), Bali, Cakici and Levy (2008), and Goyal and Santa-Clara (2003), respectively; for the equal-weighted idiosyncratic risk, the correlation values are greater than 99%,97%, and

With the evidences summarized in Table 32, it is not possible to state that idiosyncratic variance constructed from stock's market capitalization in monthly frequency forecasts market returns. In addition, the reported adjusted R^2 are mainly negative demonstrating that the variables employed are not aquedated to forecast market returns. These results indicate that for Hypothesis 6 about the importance of CSV construced according to size to predict market returns in the complete timespan, the null hypothesis cannot be rejected; hence, idiosyncratic variance based on stock's market capitalization is not relevant to forecast market returns.

Table 32 - Monthly Predictability Regressions of Market Returns using CSV by Size – Brazilian Sample

Panel A shows the results of a one-month-ahead predictive regression of the equal-weighted monthly market excess returns on the monthly lagged equal-weighted cross-sectional variance based on small (SMALL) and big (BIG) according to the stock's market capitalization for three sample periods. Panel B uses value-weighted monthly market excess returns on the monthly lagged value-weighted counterparts. The estimated regression is: $r_{t+1} = a + \beta v_t + \varepsilon_{t+1}$, where r_{t+1} represents either the equal or value-weighted market returns in period t+1, a is the constant of the model, β is the slope coefficient, v_t represents either the equal or value-weighted idiosyncratic variance based on SMALL and BIG stocks in period t, and ε_{t+1} is the model's residual. Newey-West (NW) (1987) corrected t-statistics are shown in parentheses. The sample period is from January 2000 to June 2016.

	2000-2016	2000-2007	2008-2016	2000-2016	2000-2007	2008-2016
	$CSV-SMALL^{EW}$	$CSV-SMALL^{EW}$	$CSV-SMALL^{EW}$	$CSV-BIG^{EW}$	$CSV-BIG^{EW}$	$CSV-BIG^{EW}$
Panel A: Forecasting r^{EW}						
Constant	0,008	0,019	-0,004	0,007	0,043	-0,004
NW t - statistics	(0,759)	(1,619)	(-0,247)	(1,077)	(2,173)	(-0,526)
Coefficient	-0,021	-0,049	0,019	-0,070	-2,433	0,159
NW t - statistics	(-0,165)	(-0,312)	(0,118)	(-0,294)	(-1,458)	(0,667)
$Adj.R^2$	-0,5%	-1,0%	-1,0%	-0,5%	1,2%	-0,9%
	$CSV-SMALL^{CW}$	$CSV-SMALL^{CW}$	$CSV-SMALL^{CW}$	$CSV-BIG^{CW}$	$CSV-BIG^{CW}$	$CSV-BIG^{CW}$
Panel B: Forecasting r^{CW}						
Constant	0,001	0,013	-0,006	0,005	0,023	-0,004
NW t - statistics	(0,185)	(0,777)	(-0,775)	(0,599)	(1,428)	(-0,456)
Coefficient	-0,005	-0,306	0,026	-0,478	-1,789	-0,075
NW t - statistics	(-0,028)	(-0,363)	(0,131)	(-0,571)	(-1,057)	(-0,080)
$Adj.R^2$	-0,6%	-0,9%	-1,0%	-0,3%	0,1%	-1,0%

Source: Author (2017)

98%, respeticvely. Considering the sample of small stocks, the correlation for the value-weighted measures between the CSV and those based on the based on the multifactor models (FF-3, FFC, and FF-5), Bali, Cakici and Levy (2008), and Goyal and Santa-Clara (2003) are greater than 99%, 88%, and 98%, respectively; and, for the equal-weighted version are greater than 97%, 77%, and 77%. These values show that, in general, CSV is highly correlated with other well-used approaches as in the previous case using rated firms.

4.2 Expected Idiosyncratic Volatility at the Firm-level

4.2.1 Analysis of Abnormal Returns

This subsection has the main goal of evaluating whether and how conditional idiosyncratic volatility at the firm-level is related to expected returns using a Brazilian sample. To this end, the abnormal returns generated from portfolios are investigated based on four stock's characteristics: a univariate sorting on expected idiosyncratic volatility (*Eivol*), and a bivariate sorting firstly on the stock's market capitalization, book-to-market, momentum or short term reversal followed by a sorting according to the stock's expected idiosyncratic risk.

Moreover, the Skew-GED distribution was used in the EGARCH (1,1) setup due to the presence of non-normality of financial series. Indeed, in the sample utilized here, practically 80% of stocks rejected the normality assumption given by the Jarque-Bera (1987) test.

Table 33 - Abnormal Returns on Portfolios formed on *Eivol*

Panel A reports the Fama and French (1993) three factor alphas for five value-weighted portfolios formed on *Eivol*, which represents the out-of-sample forecasting of idiosyncratic volatility of stocks controlling for the Fama and French (1993) in the mean equation and with an EGARCH (1,1) model using a Skew-GED distribution. Panels B and C represent those portfolios formed on *Eivol* using Cahart's (1997) and Fama and French's (2015) five factor models, respectively. The first portfolio (Low) consists of the bottom stocks with the lowest *Eivol*, whereas portfolio five (High) consists of the top stocks with the highest levels of *Eivol*. The column "H-L" reports the spread in alphas between portfolios High minus Low. The main specification regression is defined as: $r_t = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_t$, where r_t represents the excess portfolio return in period t, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of variables in period t, and ε_t is the model's residual. The control variables depend on the model estimated (FF-3, FFC, or FF-5). Newey-West (NW) (1987) corrected t- statistics are shown in parentheses. The sample period is from August 2004 to June 2016.

Panel A: Ranking on $Eivol$ - EGARCH + Fama and French (1993) + Skew-GED						
Variable	Low	2	3	4	High	H-L
Exchange Traded Stocks	0,001	0,001	0,005	0,002	0,007	0,006
	(0,438)	(0,215)	(2,485)	(0,936)	(2,505)	(1,980)

Panel B: Ranking on $Eivol$ - EGARCH + Carhart (1997) + Skew-GED						
Variable	Low	2	3	4	High	H-L
Exchange Traded Stocks	0,000	0,002	0,005	0,000	0,005	0,005
	(0,073)	(0,578)	(2,336)	(-0,158)	(1,821)	(1,664)

Continued

Panel C: Ranking on *Eivol* - EGARCH + Fama and French (2015) + Skew-GED						
Variable	Low	2	3	4	High	H-L
Exchange Traded Stocks	0,001	0,000	0,005	0,001	0,008	0,007
	(0,208)	(0,082)	(2,400)	(0,621)	(2,734)	(2,268)

Source: Author (2017)

One interesting point is the similarities between abnormal returns through three different models, which is corroborated by Bali *et al.* (2005) who also find very similar findings provided by these models. Also, it can be noted that the abnormal return of portfolio High, consisting of stocks at the highest level of idiosyncratic risk, shows a statistically significant positive return, except for Panel B where it is marginally significant.

More importantly, the abnormal return of portfolio H – L shows a positive association with expected returns which is in conformity with Fu’s (2009) results. Table 34 investigates the abnormal portfolio returns when controlling for one of the following characteristics: size, book-to-market, momentum, and return reversal with expected idiosyncratic volatility.

These results suggest that when controlling for an additional characteristic through a double sorting procedure, as provided in previous work related to market return, such as Fama and French’s (1992, 1993) study on the size and book-to-market, Jegadeesh and Titman’s (1993) study on momentum, and Huang *et al.*’s (2010) study on the relevance of return reversal to explain Ang *et al.*’s (2006) puzzle, the results are striking as none of the abnormal returns related to H-L is different from zero. This points out to a non-existent association between conditional idiosyncratic volatility and expected returns.

In fact, these results are also corroborated by Fink, Fink, and He (2002). When the authors use the conditional idiosyncratic volatility up to time t, as in Fu (2009), the abnormal returns of portfolio H-L is positive and significant. The same happens when controlling for other variables. However, when using the forecasted idiosyncratic volatility estimated with all the information available to traders with data up to t-1, their findings mainly document that controlling for lagged return and liquidity completely eliminates any sign of pricing quality that such strategy could earn.

Table 34 - Abnormal Returns of Portfolios formed on *Eivol* controlling for other Characteristics

Panel A reports the Fama and French (1993) three factor alphas for three value-weighted portfolios formed on *Eivol*, which represents the out-of-sample forecasting of idiosyncratic volatility of stocks controlling for the Fama and French (1993) in the mean equation and with an EGARCH (1,1) model using a Skew-GED distribution. Panels B and C represent those portfolios formed on *Eivol* using Cahart's (1997) and Fama and French's (2015) five factor models, respectively. The first portfolio (Low) consists of the bottom stocks with the lowest *Eivol*, whereas portfolio three (High) consists of the top stocks with the highest levels of *Eivol*. The column "High-Low" reports the spread in alphas between portfolios High minus Low. The main specification regression is defined as: $r_t = \alpha + \sum_{k=1}^{K} \beta_k X_{kt} + \varepsilon_t$, where r_t represents the excess portfolio return in period t, α is the constant of the model, β_k are the slope coefficients, X_{kt} refers to the set of control variables in period *t*, and ε_t is the model's residual. The control variables depend on the model estimated (FF-3, FFC, or FF-5). Newey-West (NW) (1987) corrected t- statistics are shown in parentheses. The sample period is from August 2004 to June 2016.

Panel A: Ranking on *Eivol* - EGARCH + Fama and French (1993) + Skew-GED				
Variable	Low	Medium	High	H-L
Controlling for size	0,003	0,007	-0,001	-0,004
	(1,706)	(3,145)	(-0,193)	(-1,206)
Controlling for book-to-market	0,004	0,006	0,003	-0,001
	(2,515)	(2,409)	(1,009)	(-0,561)
Controlling for momentum	0,003	0,004	0,001	-0,002
	(2,298)	(2,127)	(0,411)	(-0,703)
Controlling for return reversal	0,004	0,004	-0,000	-0,004
	(2,753)	(1,490)	(-0,127)	(-1,382)
Panel B: Ranking on *Eivol* - EGARCH + Carhart (1997) + Skew-GED				
Variable	Low	Medium	High	H-L
Controlling for size	0,003	0,005	0,002	-0,001
	(1,533)	(2,670)	(0,536)	(-0,334)
Controlling for book-to-market	0,002	0,005	0,003	0,001
	(1,191)	(2,216)	(1,348)	(0,240)
Controlling for momentum	0,002	0,004	0,002	-0,000
	(1,367)	(1,742)	(0,565)	(-0,172)
Controlling for return reversal	0,003	0,002	0,002	-0,001
	(1,753)	(0,773)	(0,714)	(-0,262)
Panel C: Ranking on *Eivol* - EGARCH + Fama and French (2015) + Skew-GED				
Variable	Low	Medium	High	H-L
Controlling for size	0,004	0,003	0,001	-0,003
	(1,722)	(1,753)	(0,258)	(-0,990)
Controlling for book-to-market	0,004	0,004	0,004	-0,000
	(2,723)	(1,500)	(1,585)	(-0,201)
Controlling for momentum	0,003	0,003	-0,001	-0,004
	(1,680)	(1,449)	(-0,333)	(-1,149)
Controlling for return reversal	0,003	0,002	-0,000	-0,003
	(2,279)	(1,104)	(-0,164)	(-1,232)

Source: Author (2017)

One negative point regarding this analysis in comparison to those presented by the United States is the number of portfolios created which mainly split stocks into quintile/deciles. Due to the reduced number of stocks traded in BM&FBOVESPA in relation to the aforementioned stock market, it was found a minimum (maximum) of 9 (40) stocks in the portfolios constructed and, hence, restricting the portfolios construction and the results presented in this subsection.

4.2.2 Fama-MacBeth Regressions

Table 35 summarizes the Fama-MacBeth regressions to analyze if expected idiosyncratic volatility is related to expected returns and controlling for a set of variables found to influence the stock returns.

From the reported results, it can be noted that the traditional equal-weighted Fama and MacBeth's (1973) regression as presented by Fu's (2009) specification (in addition to lagged returns) is opposite to what could be observed from Brazilian stocks. Forecasted idiosyncratic volatility seems to be negatively related to expected returns even when using value-weighted Fama and MacBeth's (1973) regressions, although non-significant even at the 10% level.

Also, Beta, turnover and coefficient of variation of turnover are significant in every model specification which helps to reaffirm their usefulness. The choice to include the lagged return was based on Huang *et al.* (2010), who found that this variable helps understand the puzzle found by Ang *et al.* (2006). Interestingly, as shown in Table 35, lagged return is not significant in the results.

Table 35 - Fama-MacBeth Regressions using *Eivol* and Controlling for Several Stock's Characteristics

This table presents the time series average of the slopes in cross-sectional regressions using the standard (value-weighted) Fama and MacBeth (1973) procedure in Panel A (Panel B) and estimating the Fama and French (1993), the Carhart (1997), and the Fama and French (2015) models, FF-3, FFC and FF-5, respectively, to obtain forecasted idiosyncratic volatility. The main specification regression is defined as: $r_{it} = \gamma_{0t} + \sum_{k=1}^{K} \gamma_{kt} X_{kit} + \varepsilon_{it}$, where r_{it} represents the realized return on stock i in month t, X_{kit} refers to the set of variables employed in this study, and ε_{it} is the error term. Newey-West (NW) (1987) corrected t - statistics are shown in parentheses. The sample period is from August 2004 to June 2016.

Panel A: Equal - weighted Fama - MacBeth regressions				Panel B: Cap - weighted Fama - MacBeth regressions			
Variables	FF	FFC	FF5	Variables	FF	FFC	FF5
E(IVOL)	-0,031	-0,040	-0,034	E(IVOL)	-0,036	-0,043	-0,037
	(-1,085)	(-1,387)	(-1,200)		(-1,286)	(-1,523)	(-1,368)
BETA	-0,020	-0,020	-0,020	BETA	-0,020	-0,020	-0,020
	(-2,475)	(-2,454)	(-2,464)		(-2,490)	(-2,459)	(-2,486)
LNMV	-0,001	-0,001	-0,001	LNMV	-0,001	-0,001	-0,001
	(-1,120)	(-1,196)	(-1,191)		(-1,187)	(-1,167)	(-1,178)
LNBM	0,003	0,003	0,004	LNBM	0,003	0,003	0,004
	(1,310)	(1,254)	(1,469)		(1,320)	(1,260)	(1,477)
RET(-2,-7)	0,003	0,002	0,001	RET(-2,-7)	0,003	0,002	0,001
	(0,312)	(0,196)	(0,129)		(0,312)	(0,212)	(0,135)
LN(TURN)	0,003	0,003	0,003	LN(TURN)	0,003	0,003	0,003
	(4,166)	(4,088)	(4,175)		(4,187)	(4,103)	(4,208)
LN(CVTURN)	-0,006	-0,006	-0,006	LN(CVTURN)	-0,006	-0,006	-0,006
	(-2,665)	(-2,499)	(-2,700)		(-2,706)	(-2,536)	(-2,720)
REV	-0,021	-0,020	-0,023	REV	-0,020	-0,019	-0,022
	(-1,485)	(-1,431)	(-1,614)		(-1,433)	(-1,369)	(-1,566)
$Adj.R^2$	12,4%	12,4%	12,3%	$Adj.R^2$	12,3%	12,4%	12,3%

Source: Author (2017)

Overall, these findings suggest that there is no robust association between expected idiosyncratic volatility and expected returns when considering an adequate number of iterations during the estimation procedure, using a non-normal distribution to accommodate the characteristics of the financial series, and making use of only the information available to traders at that moment, i.e. without a forward-looking return observation. Hence, regarding Hypothesis 7 about a significant association between expected idiosyncratic volatility at the firm-level and expected returns, the analyses indicated that the null hypothesis cannot be rejected in this study.

Lastly, Table 36 summarizes the main results found in this study.

Table 36 – Summary of Results

Relevance of aggregate idiosyncratic risk in predicting market returns - Hypothesis 1
Results show, based on the whole sample, CSV^{EW} and CSV^{CW} play a central role in predicting equal – and value-weighted market returns, respectively, only for the United States and it is independent of the frequency of data used.
Influence of investor sentiment on the relation between aggregate idiosyncratic risk in predicting market returns – Hypothesis 2
The evidence presented indicates that he null hypothesis that idiosyncratic variance does not matter in the presence of market variance and investor sentiment can be rejected considering only the equal-weighted measure for the United States sample, while for the Brazilian case it is rejected in both weighting schemes.
Role of CSV in predicting market returns in the presence of higher cross-sectional moments – Hypothesis 3
The analyses for the United States sample in whole period show that with the exception of CSV^{CW} to predict daily value-weighted market returns, all the remaining estimations indicate that daily or monthly idiosyncratic variance matters in the presence of higher cross-sectional moments. For the Brazilian sample, only CSV^{EW} is relevant to forecast daily equal-weighted market returns with higher cross-sectional moments in the specification, i.e, these results point to the lack of power of idiosyncratic variance in the emerging country studied.
Forecasting power of cross-sectional variance for size, value, momentum, investment, and profitability factors – Hypothesis 4
For the United States sample, there is limited evidence of the importance of idiosyncratic variance in predicting market returns and controlling for the corresponding market variance weighting-scheme as the estimations indicate that only CSV^{CW} is able to predict the profitability factor (RMW). For the Brazilian case, neither CSV^{EW} nor CSV^{CW} are useful variables.

Continued

Usefulness of a measure of aggregate idiosyncratic variance based on credit ratings for the U.S. and Brazilian cases – Hypothesis 5
For the United States sample in the whole timespan, idiosyncratic variance based on NIG, IG, and all rated firms do not forecast daily equal- or value-weighted market returns, but monthly CSV^{CW} according to each of these three categories matter. For the Brazilian case, with exception of the relevance of daily $CSV-RAT^{CW}$ to forecast daily value-weighted market returns, none of the other models in daily or monthly frequency suggest that these measures matter.
Usefulness of a measure of aggregate idiosyncratic variance based on size for the Brazilian stock market in predicting market returns – Hypothesis 6
Using the complete timespan, the null hypothesis cannot be rejected; hence, idiosyncratic variance based on stock's market capitalization is not relevant to forecast market returns.
Association between expected idiosyncratic volatility and expected returns using a Brazilian sample of stocks – Hypothesis 7
There is no robust association between expected idiosyncratic volatility and expected returns when considering an adequate number of iterations during the estimation procedure, using a non-normal distribution to accommodate the characteristics of the financial series, and without including the contemporaneous return.

Source: Author (2017)

5 CONCLUSIONS

It is well known that the risk of a stock is split into two components: the systematic part that is linked to the portion which cannot be vanished away even if investors try to have a well-diversified portfolio, and the unsystematic risk which represents the one based on the firm-specific risk. Modern Portfolio Theory assumes that investors are rational and diversify away the firm-specific risk. The CAPM model is based on the work of Markowitz (1952, 1959) and its underlying assumptions reflect that investors demand a risk premia according to the exposure to the correlation of the portfolio returns with the market returns, evidenced on the systematic part.

However, under-diversification theories, as in Merton (1987), posit that investors do not fully diversify away their exposure to the unsystematic risk, also called idiosyncratic risk and commonly measured by the idiosyncratic volatility, and indicates as one of the reasons the existence of transaction costs and limited information. The fact that stocks may have short-sell constraints is also a reason for this phenomenon.

If investors do not diversify their portfolios, idiosyncratic volatility may be a risk factor that affects the investor portfolio and if this is true then investors would require a risk premia for dealing with it. Additionally, it is commonly understood that higher risk should generate higher return, then if it is a priced factor, it is expected to have a positive relationship with future returns. Indeed, theories such as Merton's (1987) and Malkiel and Xu's (2002) refer to a positive relation between the unsystematic risk and expected returns.

In contrast, there is empirical evidence pointing out that idiosyncratic volatility at the aggregate level is positively related to market returns, such as Goyal and Santa-Clara's and Garcia, Mantilla-García, and Martellini's (2014), a negative relation as demonstrated in Guo and Savickas (2006, 2009) and corroborated for some countries in Angelidis (2010) and a non-significant relation as presented by Bali *et al.* (2005) and Wei and Zhang (2005).

At the firm-level, there is also contradictory evidence. Ang *et al.*'s (2006) study of the U.S stock market and Ang *et al.* (2009), including 23 countries, report a negative relation between realized (lagged) idiosyncratic volatility and expected stock returns while Fu (2009) recovers a positive relation between forecasted idiosyncratic volatility and expected returns applying the Fama and French (1993) model in the mean equation and an EGARCH volatility model. The author documents that realized idiosyncratic volatility, as used in Ang *et al.*

(2006, 2009) is not a correct proxy for measuring firm-specific risk due to its time-varying behavior over time.

Meanwhile, although Fu (2009) suggests obtaining the firm-specific risk from a conditional volatility perspective, there is a forward-looking bias in his estimations with the inclusion of the return observation in month *t* to predict conditional firm-specific risk also in month *t*. In contrast, Guo, Kassa, and Ferguson (2014) find that there is no relation between expected returns and conditional idiosyncratic volatility if the one-forward looking observation is removed. Fink, Fink, and He (2012) support their findings, arguing that the association between them is statically insignificant.

In this sense, this work focused on the study of idiosyncratic risk at the aggregate and firm-level. Relying on the cross-sectional variance of return as given in Garcia, Mantilla-García, and Martellini (2014) as a proxy to obtain idiosyncratic variance in a daily and monthly frequency this study verified its predictive power for market returns and risk factors omitted from the CAPM model as well as performed different sets of regressions using CSV alone or adding market variance, investor sentiment, aggregate dividend yield and measures of illiquidity. Equal- and value-weighted market returns were used to this end.

A CSV measure according to the firm's credit rating was built to verify if the idiosyncratic volatility of firms according to their financial distress as proxied by their ratings could be useful to predict market returns for the United States and Brazilian stock market, but constraints on the number of Brazilian firms rated brought limitation to this analysis as it reduces the number of stocks employed to the consctruction of CSV. Trying to overcome this issue, the stock's market capitalization was considered as a control characteristic to obtain idiosyncratic variance, as suggested by Brown and Ferreira (2004) and followed by Angelidis and Tessaromatis (2008) for the United Kingdom stock market.

The results were as follows: for the United States sample of stocks, CSV remains a relevant predictor of market returns, but the power of CSV^{EW} is more robust to different specifications and periods than its value-weighted version, CSV^{CW}. In contrast, the same cannot be said for Brazilian stocks even with the analysis extended to subperiods considering both weighting schemes. The findings do not support the hypothesis 1 about the importance of aggregate idiosyncratic risk to forecast market returns using the Brazilian stock market.

Hypothesis 2 concerning the investor sentiment shows that for the United States investor sentiment is significant only when used together with equal-weighted idiosyncratic variance, while the latter remains its predictability power. However, using a sample of Brazilian stocks it was found that investor sentiment and equal- and value-weighted idiosyncratic variance are not relevant to forecast market returns.

In relation to hypothesis 3 concerning the higher cross-sectional moments, it can be said that for the United States the cross-sectional skewness (CSS) shows to be more important than cross-sectional kurtosis (CSK). More importantly, in daily regressions and considering the extended sample, CSV^{EW} mantains its statistical power in the presence of cross-sectional higher moments while CSV^{CW} loses power in the predictive regressions. In a monthly frequency $CSV^{EW}$and CSV^{CW} are relevant to forecast market returns even in the presence of CSS and CSK.

Using the Brazilian sample, this research found that CSK is more valuable than CSS and, in general, the results suggest that cross-sectional variance is not important in the full specification model as only the CSV^{EW} in a daily frequency to predict equal-weighted market returns matters taking into account two different weighting schemes to consctruct idiosyncratic variance (equal- and value-weighted market returns) and two frequency of data (daily and monthly). Another verification done in this thesis was to evaluate the power of idiosyncratic variance to predict risk factors. In this sense, hypothesis 4 could be rejected as for both markets only limited evidence was found about their forecasting ability.

Additionally, the investigation of hypothesis 5 about the importance of CSV based on investment grade, non-investment grade and all rated firms for the complete timespan suggests that the value-weighted measures of idiosyncratic variance were important to predict market returns, but only on a monthly basis. The inferences for Brazilian stocks show limited evidence where daily $CSV - RAT^{CW}$ was the only case which was shown to be able to predict value-weighted market returns. In summary, neither CSV by levels of ratings, where due to limitations was split only into non-investment grade and all rated firms, nor by size, small versus big, help to consistently forecast market returns (hypothesis 6).

Moreover, an analysis at the firm-level through three different models means equation specifications (Fama and French's 1993, Carhart', 1997, and Fama and French's 2015 models) and the EGARCH (1,1) volatility model, using data up to time t-1 or, in other words, using only the information available to traders, adopting the Skew-GED as a more flexible distribution to account for deviations from normality in stock returns and a setting with 2.000 iterations, the Fama and MacBeth (1973) cross-sectional regressions corroborate Fink, Fink, and He’s (2012) and Guo, Kassa, and Ferguson’s (2014) statement that the relation between expected returns and forecasted idiosyncratic volatility is non-existent (hypothesis 7).

Overall, these findings implicate that idiosyncratic variance, proxied by cross-sectional variance of returns, and expected idiosyncratic volatility does not play a relevant role from the perspective of an emergent country whereas the main body of empirical tests are conducted on the United States and other developed countries around the world.

Five main possibilities for future research are also given here. First, the use of quarterly data to identify whether using a different frequency of data alters the results presented here. Second, the comparison of the measure cross-sectional variance as employed here with that based on portfolios followed in Angelidis, Sakkas, and Tessaromatis (2015) can contribute to the literature of idiosyncratic variance (volatility) using Brazilian data.

Third, it would be very interesting to understand the role of institutional investor (XU; MALKIEL, 2003), retail trading (BRANDT *et al*., 2010) and investor sentiment (FINK *et al.*, 2010) behaviors in affecting (or not) the trend in idiosyncratic volatility from an emergent country view. Fourth, it would be also relevant to investigate the behavior of investor sentiment on the conditional idiosyncratic volatility as presented here and to verify if the relation differs when controlling for high-sentiment and low-sentiment periods and expanding this analysis to other countries in Latin America. Finally, understanding the behavior of idiosyncratic variance in predicting business cycles compared to developed countries would help to fill a gap in this literature applied in local market.

REFERENCES

ACHARYA, V. V.; PEDERSEN, L. H. Asset pricing with liquidity risk. **Journal of Financial Economics**, v. 77, p. 375 – 410, 2005.

ALTMAN, E. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. **Journal of Finance**, v. 23, n. 4, p. 589 – 609, 1968.

AMIHUD, Y. Illiquidity and stock returns. Cross-section and time-series effects. **Journal of Financial Markets**, v. 5, n. 1, p. 31 – 56, 2002.

ANG, A.; HODRICK, R. J.; XING, Y.; ZHANG, X. The cross-section of volatility and expected returns. **Journal of Finance**, v. 61, n.1, p. 259 – 299, 2006.

________; ________; ________; ________. High idiosyncratic volatility and low returns: International and further U.S. evidence. **Journal of Financial Economics**, v. 91, p. 1 – 23, 2009.

ANGELIDIS, T.; TESSAROMATIS, N. Idiosyncratic volatility and equity returns: UK evidence. **International Review of Financial Analysis**, v. 17, p. 539 – 556, 2008.

________. Idiosyncratic risk in emerging markets. **The Financial Review**, v. 45, p.1053 – 1078, 2010.

________; SAKKAS, A.; TESSAROMATIS, N. Stock market dispersion, the business cycle and expected factor returns. **Journal of Banking & Finance**, v. 59, p. 265 – 279, 2015.

ASHBAUGH-SKAIFE, H.; COLLINS, D. W.; LAFOND, R. The effects of corporate governance on firm's credit ratings. **Journal of Accounting and Economics**, v. 42, p. 203 – 243, 2006.

AVRAMOV, D.; CHORDIA, T.; JOSTOVA, G.; PHILIPOV, A. Momentum and credit rating. **Journal of Finance**, v. 62, n. 5, p. 2503 – 2520, 2007.

________; ________; ________; ________. Credit ratings and the cross-section of stock returns. **Journal of Financial Markets**, v. 12, p. 469 – 499, 2009.

________; ________; ________; ________. Anomalies and financial distress. **Journal of Financial Economics**, v. 108, n. 1, p. 139 – 159, 2013.

BAE, K-H.; LIM,C.; WEI, K. C. J. Corporate governance and conditional skewness in the world's stock markets. **The Journal of Business**, v. 79, n. 6, p. 2999 - 3028, 2006.

BAI, X.; RUSSELL, J. R.; TIAO, G. C. Kurtosis of GARCH and stochastic volatility models with non-normal innovations. **Journal of Econometrics**, v. 114, p. 349 - 360, 2003.

BAKER, M.; WURGLER, J. Investor sentiment and the cross-section of stock returns. **Journal of Finance**, v. 61, n. 4, p. 1645 – 1680, 2006.

BANZ, R. W. The relationship between return and market value of common stocks. **Journal of Financial Economics**, v. 9, p. 3 - 18, 1981.

BALI, T. G.; CAKICI, N.; YAN, X.; ZHANG, Z. Does idiosyncratic risk really matter? **Journal of Finance**, v. 60, n. 2, p.905 – 929, 2005.

________; ________; Idiosyncratic volatility and the cross section of expected returns. **Journal of Financial and Quantitative Analysis**, v. 43, n. 1, p. 29 – 58, 2008.

________; ________; LEVY, H. A model-independent measure of aggregate risk. **Journal of Empirical Finance**, v. 15, p. 878 – 896, 2008.

________; ________; WHITELAW, R. F. Maxing out: Stocks as lotteries and the cross-section of expected returns. **Journal of Financial Economics**, v. 99, p. 427 – 446, 2011.

BARBERIS, N.; HUANG, M. Mental accounting, loss aversion, and individual stock returns. **Journal of Finance**, v. 56, n. 4, p. 1247 – 1292, 2001.

BASU, S. The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. **Journal of Financial Economics**, v. 12, p. 129 – 156, 1983.

BEKAERT, G.; HODRICK, R. J.; ZHANG, X. Aggregate idiosyncratic volatility. **Journal of Financial and Quantitative Analysis**, v. 47, n. 6, p. 1155 – 1185, 2012.

BHANDARI, L. C. Debt/Equity ratio and expected common stock returns: Empirical evidence. **Journal of Finance**, v. 43, n. 2, p. 507 – 528, 1988.

BLACK, F. Capital market equilibrium with restricted borrowing. **The Journal of Business**, v. 45, n. 3, p. 444 - 455, 1972.

________. Studies of stock price volatility changes. Proceedings of the 1976 meetings of the Business and Economics Statistics Section, **American Statisical Association**, p. 177 - 181, 1976.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, p. 307 – 327, 1986.

BOWLEY, A. L. **Elements of Statistics**, v. 2. London: P. S. King and Son, 1920.

BOYER, B.; MITTON, T.; VORKINK, K. Expected idiosyncratic skewness. **Review of Financial Studies**, v. 23, n. 1, p. 169 – 202, 2010.

BRANDT, M. W.; BRAV, A.; GRAHAM, J. R.; KUMAR, A. The idiosyncratic volatility puzzle: Time trend or speculative episodes? **Review of Financial Studies**, v. 23, n. 2, p. 863 – 899, 2010.

BREUSCH, T.; PAGAN, A. A Simple Test of Heteroskedasticity and Random Coefficient Variation. **Econometrica**, v.47, p. 1287 - 1294, 1979.

BROCKMAN, P.; SCHUTTE, M. G.; YU, W. Is idiosyncratic risk priced? The international evidence. **Working paper**, Lehigh University, p. 1- 52, 2009.

BROOKS, C.; BURKE, S. P.; HERAVI, S.; PERSAND, G. Autoregressive conditional kurtosis. **Journal of Financial Econometrics**, v. 3, n. 3, p. 399 – 421, 2005.

BROWN, D. P.; FERREIRA, M. A. Information in the idiosyncratic volatility of small firms. **Working paper**, EFA 2004 Maastricht Meeting Paper, p. 1-58, 2004.

BROWN, G. W.; CLIFF, M. T. Investor sentiment and asset valuation. **Journal of Business**, v. 78, n. 2, p. 405 – 440, 2005.

CAMPBELL, J. Y.; LETTAU, M.; MALKIEL, B. G.; XU, Y. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. **Journal of Finance**, v. 56, n. 1, p. 1 – 43, 2001.

CAO, C.; SIMIN, T.; ZHAO, J. Can growth options explain the trend in idiosyncratic risk? **Review of Financial Studies**, v. 21, n. 6, p. 2599 – 2633, 2008.

CAO, H.; WANG, T.; ZHANG, H. Model uncertainty, limited market participation and asset prices. **Review of Financial Studies**, v. 18, n. 4, p. 1219 – 1251, 2005.

CARHART, M. M. On persistence in mutual fund performance. **Journal of Finance**, v. 52, n. 1, p. 57 – 82, 1997.

CHAN, L. K. C.; HAMAO, Y.; LAKONISHOK, J. Fundamentals and stock returns in Japan. **Journal of Finance**, v. 46, n. 5, p. 1739 – 1764, 1991.

CHEN, J.; HONG, H.; STEIN, J. C.; Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. **Journal of Financial Economics**, v. 61, p. 345 - 381, 2001.

CHORDIA, T. SUBRAHMANYAM, A.; ANSHUMAN, V. Trading activity and expected stock returns. **Journal of Financial Economics**, v. 59, n. 1, p. 3 – 32, 2001.

CHRISTIE, A. A. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. **Journal of Financial Economics**, v. 10, p. 407 - 432, 1982.

CHUA, C. T.; GOH, J.; ZHANG, Z. Expected volatility, unexpected volatility, and the cross-section of stock returns. **Journal of Financial Research**, v. 33, n. 2, p. 103 – 123, 2010.

CONT, R. Empirical properties of asset returns: Stylized facts and statistical issues. **Quantitative Finance**, v. 1, p. 223 - 236, 2001.

COSTA, H. C.; MAZZEU, J. H. G.; COSTA JUNIOR, N. C. A. O comportamento dos componentes da volatilidade das ações no Brasil. **Revista Brasileira de Finanças**, v. 14, n. 2, p. 225 – 268, 2016.

CRESWELL, J. W. **Projeto de pesquisa**: Método qualitativo, quantitativo e misto. 2 ed. Porto Alegre: Artmed, 2007.

CROW, E. L.; SIDDIQUI, M. M. Robust estimation of location**. Journal of the American Statistical Association**, v. 62, n. 318, p. 353 – 389, 1967.

DAMODARAN, A. Economic events, information structure, and the return-generating process. **The Journal of Financial and Quantitative Analysis**, v. 20, n. 4, p. 423 - 434, 1985.

DAVIDSON, R.; MACKINNON, J. G. **Estimation and inference in Econometrics**. New York: Oxford University Press, 1993.

DITTMAR, R. F. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. **Journal of Finance**, v. 57, n. 1, p. 369 - 403, 2002.

DOUGLAS, G. W. Risk in the equity markets: An empirical appraisal of market efficiency. **Yale Economic Essays**, v. 9, p. 3 – 45, 1969.

ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, v. 50, n. 4, p. 987 – 1007, 1982.

________; MUSTAFA, C. Implied ARCH models from option prices. **Journal of Econometrics**, v. 52, n. 1, p. 289 – 311, 1992.

FAMA, E. F. The behavior of stock market prices. **Journal of Business**, v. 38, n. 1, p. 34 – 105, 1965.

________. Risk, return and equilibrium: Some clarifying comments. **Journal of Finance**, v. 23, n. 1, p. 29 - 40, 1968.

________; MACBETH, J. D. Risk, return, and equilibrium: Empirical tests. **The Journal of Political Economy**, v. 81, n. 3, p. 607 - 636, 1973.

________; FRENCH, K. R. The cross-section of expected stock returns. **Journal of Finance**, v. 47, n. 2, p. 427 - 465, 1992.

________; ________. Common risk factors in the returns on stocks and bonds. **Journal of Financial Economics**, v. 33, p. 3 – 56, 1993.

________;________. The capital asset pricing model: Theory and evidence. **The Journal of Economic Perspectives**, v. 18, n. 3, p. 25 - 46, 2004.

________;________. A five-factor asset pricing model. **Journal of Financial Economics**, v. 116, n. 1, p. 1 - 22, 2015.

FANG, H.; LAI, T. Y. Co-kurtosis and capital asset pricing. **Financial Review**, v. 32, n. 2, p. 293 – 307, 1997.

FERNANDEZ, C.; STEEL, M. F. J. On bayesian modeling of fat tails and skewness. **Jounal of the American Statistical Association**, v. 93, n. 441, p. 359 – 371, 1998.

FEUNOU, B.; JAHAN-PARVAR, M. R.; TÉDONGAP, R. Which parametric model for conditional skewness? **The European Journal of Finance**, p. 1 - 35, 2014.

FINK, J.; FINK, K. E.; GRULLON, G.; WESTON, J. P. What drove the increase in idiosyncratic volatility during the internet boom? **Journal of Financial and Quantitative Analysis**, v. 45, n. 5, p. 1253 – 1278, 2010.

________; ________; HE, H. Expected idiosyncratic volatility measures and expected returns. **Financial Management**, v. 41, n. 3, p. 519 - 553, 2012.

FRENCH, K.; SCHWERT, G.; STAMBAUGH, R. Expected stock returns and volatility. **Journal of Financial Economics**, v.19, n. 1, p. 3 – 29, 1987.

FU, F. Idiosyncratic risk and the cross-section of expected stock returns. **Journal of Financial Economics**, v. 91, p. 24 – 37, 2009.

________. On the robustness of the positive relation between expected idiosyncratic volatility and return. **Working paper**, Singapore Management University p. 1 – 6, 2010.

GALAGEDERA, D. U. A. A review of capital asset pricing models. **Managerial Finance**, v. 33, n. 10, p. 821 - 832, 2007.

GALDI, F. C.; SECURATO, J. R. O risco idiossincrático é relevante no mercado brasileiro? **Revista Brasileira de Finanças**, v. 5, n. 1, p. 41 – 58, 2007.

GARCIA, R.; MANTILLA- GARCÍA, D.; MARTELLINI, L. A model-free measure of aggregate idiosyncratic volatility and the prediction of market returns. **Journal of Financial and Quantitative Analysis**, v. 49, nos. 5/6, p. 1133 – 1165, 2014.

GHYSELS, E.; SANTA-CLARA, P.; VALKANOV, R. There is a risk-return trade-off after all. **Journal of Financial Economics**, v. 76, p. 509 – 548, 2005.

GIBBONS, M. R.; ROSS, S. A.; SHANKEN, J. A test of the efficiency of a given portfolio. **Econometrica**, v. 57, n. 5, p. 1121 - 1152, 1989.

GOETZMANN, W. N.; KUMAR, A. Equity portfolio diversification. **Review of Finance**, v. 12, p. 433 – 463, 2008.

GOYAL, A.; SANTA-CLARA, P. Idiosyncratic risk matters! **Journal of Finance**, v. 58, n. 3, p. 975 – 1007, 2003.

GREENE, W. H. **Econometric analysis**. 5th ed. New York: Prentice Hall, 2002.

GROENEVELD, R. A.; MEEDEN, G. Measuring skewness and kurtosis. Journal of the Royal Statistical Society (**The Statistician)**, v. 33, n. 4, p. 391 – 399, 1984.

GUO, H. Limited stock market participation and asset prices in a dynamic economy. **Journal of Financial and Quantitave Analysis**, v. 39, n. 3, p. 495 – 516, 2004.

________; SAVICKAS, R. Idiosyncratic volatility, stock market volatility, and Expected Stock Returns. **Journal of Business & Economic Statistics**, v. 24, n. 1, p. 43 – 56, 2006.

________; ________. Average idiosyncratic volatility in G7 countries. **Review of Financial Studies**, v. 21, p. 1259 – 1296, 2008.

________; KASSA, H.; FERGUSON, M. F. On the relation between EGARCH idiosyncratic volatility and expected stock returns. **Journal of Financial and Quantitative Analysis**, v. 49, n. 1, p. 271 – 296, 2014.

HAN, B.; KUMAR, A. Speculative retail trading and asset prices. **Journal of Financial and Quantitative Analysis**, v. 48, n. 2, p. 377 – 404, 2013.

HARVEY, C. R.; SIDDIQUE, A. Autoregressive conditional skewness. **Journal of Financial and Quantitative Analysis**, v. 34, n. 4, p. 465 – 487, 1999.

HASBROUCK, J. Trading costs and returns for US equities: The evidence from daily data. **Working paper**, Stern School of Business, p. 1 – 42, 2005.

HINKLEY, D. V. On power transformations to symmetry. **Biometrika**, v. 62, n. 1, p. 101 – 111, 1975.

HUANG, W.; LIU, Q.; RHEE, S. G.; ZHANG, L. Return reversals, idiosyncratic risk, and expected returns. **Review of Financial Studies**, v. 23, n. 1, p. 147 – 168, 2010.

JARQUE, C. M.; BERA, A. K. A test for normality of observations and regression residuals. **International Statistical Review**, v. 55, n. 2, p. 163 – 172, 1987.

JEGADEESH, N.; TITMAN, S. Returns to buying winners and selling losers: Implications for stock market efficiency. **Journal of Finance**, v. 48, n. 1, p. 65 - 91, 1993.

JIANG, X.; LEE, B-S. The dynamic relation between returns and idiosyncratic volatility. **Financial Management**, v. 35, n. 2, p. 43 – 65, 2006.

JONDEAU, E.; ROCKINGER, M. Optimal portolio allocation under higher moments. **European Financial Management**, v. 12, n. 1, p. 29 - 55, 2006.

KEARNEY, C.; POTI, V. Have European stocks become more volatile? An empirical investigation of idiosyncratic and market risk in the Euro area. **European Financial Management**, v. 14, n. 3, p. 419 – 444, 2008.

KIM, T-W.; WHITE, H. On more robust estimation of skewness and kurtosis. **Financial Research Letters**, v. 1, p. 56 – 73, 2004.

KRAUS, A.; LITZENBERGER, R. H. Skewness preference and the valuation of risk assets. **Journal of Finance**, v. 31, n. 4, p. 1085 – 1100, 1976.

LEHMANN, B. N. Residual risk revisited. **Journal of Econometrics**, v. 45, p. 71 – 97, 1990.

LEMMON, M.; PORTNIAGUINA, E. Consumer confidence and asset prices: Some empirical evidence. **Review of Financial Studies**, v. 19, n. 4, p. 1499 – 1529, 2006.

LETTAU, M.; LUDVIGSON, S. Consumption, aggregate wealth, and expected stock returns. **Journal of Finance**, v. 56, n. 3, p. 815 – 849, 2001.

LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The Review of Economics and Statistics**, v. 47, n. 1, p. 13 – 37, 1965a.

________. Security prices, risk, and maximal gains from diversification. **Journal of Finance**, v. 20, n. 4, p. 587 – 615, 1965b.

LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, v. 65, p. 297 – 303, 1978.

MALKIEL, B. G.; XU, Y. Risk and return revisited. **Journal of Portfolio Management**, v. 23, n. 3, p. 9 – 14, 1997.

________; ________. Idiosyncratic risk and security returns. **Working paper**, University of Texas at Dallas, p. 1 – 48, 2002.

MANDELBROT, B. The variation of certain speculative prices. **The Journal of Business**, vol. 36, n.4, p. 394 – 419, 1963.

MARKOWITZ, H. Portfolio selection. **Journal of Finance**, v. 7, n. 1, p. 77 – 91, 1952.

________. **Portfolio selection**: Efficient diversification of investments. USA: John Wiley & Sons, 1959.

MENDONÇA, F. P.; KLOTZLE, M. C.; PINTO, A. C. F.; MONTEZANO, R. M. S. The relationship between idiosyncratic risk and returns in the Brazilian stock market. **Revista Contabilidade e Finanças**, v. 23, n. 60, p. 246 – 257, 2012.

MERTON, R. C. An intertemporal capital asset pricing model. **Econometrica**, v. 41, p. 867 - 887, 1973.

________. A simple model of capital market equilibrium with incomplete information. **Journal of Finance**, v. 42, n. 3, p. 483 – 510, 1987.

MILLER, M.; SCHOLES, M. Rates and returns in relation to risk: A re-examination of some recent findings. In: JENSEN, M. C. (Org.). **Studies in the theory of capital markets**. New York: Praeger, 1972, p. 47 – 78.

MITTON, T.; VORKINK, K. Equilibrium undersification and the preference for skewness. **Review of Financial Studies**, v. 20, n. 4, p. 1255 – 1288, 2007.

MOSSIN, J. Equilibrium in a capital asset market. **Econometrica**, v. 34, n. 4, p. 768 – 783, 1966.

MUSSA, A.; ROGERS, P.; SECURATO, J. R. Modelos de retornos esperados no mercado brasileiro: testes empíricos utilizando metodologia preditiva. **Revista de Ciências da Administração**, v. 11, n. 23, p. 192 – 216, 2009.

______; FAMÁ, R.; SANTOS, J. O. A adição do fator de risco momento ao modelo de precificação de ativos dos três fatores de Fama & French aplicado ao mercado acionário brasileiro. **Revista de Gestão**, v. 19, n. 3, p. 431 – 447, 2012.

NELSON, D. B. Conditional Heteroskedasticity in asset returns: a new approach. **Econometrica**, v. 59, n.2, p. 347 – 370,1991.

NEWEY, W. K; WEST, K. D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. **Econometrica**, v. 55, n. 3, p. 703 – 708, 1987.

NOVY-MARX, R. The other side of value: The gross profitability premium. **Journal of Financial Economics**, v. 108, p. 1 – 28, 2013.

PAGAN, A. R.; SCHWERT, G. W. Testing for covariance stationarity in stock market data. **Economic Letters**, v. 33, n. 2, p. 165 – 170, 1990.

PÁSTOR, L.; VERONESI, P. Stock valuation and learning about profitability. **Journal of Finance**, v. 58, n. 5, p. 1749 – 1789, 2003.

PEIRÓ, A. Skewness in financial data. **Journal of Banking and Finance**, v. 23, p. 847 - 862, 1999.

PETERSON, D. R.; SMEDEMA, A. R. The return impact of realized and expected idiosyncratic volatility. **Journal of Banking & Finance**, v. 35, p. 2547 – 2558, 2011.

ROLL, R. A critique of the asset pricing theory's tests. Parte I: On past and potential testability of the theory. **Journal of Financial Economics**, v. 4, p. 129 – 176, 1977.

ROSENBERG, B.; REID, K.; LANSTEIN, R. Persuasive evidence of market inefficiency. **The Journal of Portfolio Management**, v. 11, n. 3, p. 9 - 16, 1985.

ROSS, S. A. The arbitrage theory of capital asset pricing. **Journal of Economic Theory**, v. 13, p. 341 – 360, 1976.

SCHMELING, M. Investor sentiment and stock returns: Some international evidence. **Journal of Empirical Finance**, v. 16, p. 394 – 408, 2009.

SEHGAL, S.; GARG, V. Cross sectional moments and portfolio returns: Evidence for select emerging markets. **IIMB Management Review**, v. 28, p. 147 – 159, 2016.

SHARPE, W. F. Capital asset prices: A theory of market equilibrium under conditions of risk. **Journal of Finance**, v. 19, n. 3, p. 425 – 442, 1964.

SHIH, Y-C.; CHEN, S-S.; LEE, C-F.; CHEN, P-J. The evolution o capital asset pricing models. **Review of Quantitative Finance and Accounting**, v. 42, p. 415 - 448, 2014.

SPIEGEL, M. I.; WANG, X. Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk. **Working paper**, Yale University, p. 1 – 49, 2005.

STANDARD & POOR'S. **S&P Global Ratings Definitions**. 2016. Disponível em: <https://www.standardandpoors.com/en_EU/delegate/getPDF?articleId=1824955&type=COMMENTS&subType=REGULATORY>. Acesso em: 12 dez. 2016.

STATMAN, M. How many stocks make a diversified portfolio? **Journal of Financial and Quantitative Analysis**, v. 22, n. 3, p. 353 – 363, 1987.

STATTMAN, D. Book values and stock returns. **The Chicago MBA: A Journal of Selected Papers**, v. 4, p. 25 – 45, 1980.

STIVERS, C.; SUN, L. Cross-sectional return dispersion and time variation in value and momentum premiums. **Journal of Financial and Quantitative Analysis**, v. 45, n. 4, p.987 – 1014, 2010.

SU, J-B.; LEE, M-C.; CHIU, C-L. Why does skewness and the fat tail effect influence value-at-risk estimates? Evidence from alternative capital markets. **International Review of Economics and Finance**, v. 31, p. 59 - 85, 2014.

TANG, T. T. Information asymmetry and firm's credit market access: Evidence from Moody's credit rating format refinement. **Journal of Financial Economics**, v. 93, n. 2, p. 325 – 351, 2009.

TITMAN, S.; WEI, K. C. J.; XIE, F. Capital investments and stock returns. **Journal of Financial and Quantitative Analysis**, v. 39, n. 4, p. 677 – 700, 2004.

VERHOEVEN, P.; MCALEER, M. Fat tails and asymmetry in financial volatility models. **Mathematics and Computers in Simulation**, v. 64, p. 351 - 361, 2004.

XU, Y.; MALKIEL, B. G. Investigating the behavior of idiosyncratic volatility**. Journal of Business**, v. 76, n. 4, p. 613 – 645, 2003.

WAN, C.; XIAO, Z. Idiosyncratic volatility, expected windfall, and the cross-section of stock returns. **Advances in Econometrics**, v. 33, p. 713 – 749, 2014.

WEI, S. X.; ZHANG, C. Idiosyncratic risk does not matter: A re-examination of the relationship between average returns and average volatilities. **Journal of Banking & Finance**, v. 29, p. 603 – 621, 2005.