YAŞAR UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
BUSINESS ADMINISTRATION
DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

THE EFFECT OF INTERNET SEARCH VOLUME (ISV) ON STOCK RETURN VOLATILITY

SEMEN SON

Supervisor: Assist. Prof. Dr. R. Serkan Albayrak
Co-Supervisor: Prof. Dr. Pınar Evrim Mandacı

İzmir, 2014

TUTANAK

T.C.

YAŞAR ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ DOKTORA TEZİ JÜRİ SINAV TUTANAĞI

ÖĞRENCİNİN	
Adı, Soyadı	: Semen SON
Öğrenci No	: 1130000 9002
Anabilim Dalı	: İşletme
Programı	: Doktora
Tez Sınav Tarihi	: 17/01/2014 Sınav Saati : 14:00

Tezin Başlığı: The Effect of ISV on Stock Return Volatility

Adayın kişisel çalışmasına dayanan tezini .... dakikalık süre içinde savunmasından sonra jüri üyelerince gerek tez konusu gerekse tezin dayanağı olan anabilim dallarından sorulan sorulara verdiği cevaplar değerlendirilerek, tezin ;

☑ BAŞARILI olduğuna (S) ☑ OY BİRLİĞİ
1 ☐ DÜZELTME gerekliliğine (I) ile karar verilmiştir.
2 ☐ BAŞARISIZ sayılmasına (F) ☐ OY ÇOKLUĞU

3 ☐ Jüri toplanamadığı için sınav yapılamamıştır.
4 ☐ Öğrenci sınava gelmemiştir.

☑ Başarılı ☐ Düzeltme ☐ Başarısız	☑ Başarılı ☐ Düzeltme ☐ Başarısız	☑ Başarılı ☐ Düzeltme ☐ Başarısız	☑ Başarılı ☐ Düzeltme ☐ Başarısız	☑ Başarılı ☐ Düzeltme ☐ Başarısız
Üye: Prof. Dr. Pınar Evrim Mandacı	Üye: Prof. Dr. Hasan F. Baklacı	Üye: Yrd.Doç.Dr. R. Serkan Albayrak	Üye: Çağrı Bulut	Üye: Yrd.Doç.Dr. F. Dilvin Taşkın
İmza:	İmza:	İmza:	İmza:	İmza:

1) *Bu halde adaya 6 ay süre verilmiştir.*
2) *Bu halde öğrencinin kaydı silineceğinden jüri red kararı hakkında en az bir sayfalık bir gerekçe yazarak sınav tutanağına eklemelidir.*
3) *Bu halde sınav için yeni bir tarih belirlenir.*
4) *Bu halde varsa öğrencinin mazeret belgesi Enstitü Yönetim Kurulunda görüşülür.Öğrencinin geçerli mazereti olmaması halinde Enstitü Yönetim Kurulu kararı ile ilişiği kesilir.Mazereti geçerli sayıldığında yeni bir sınav tarihi belirlenir.*

ACKNOWLEDGEMENTS

Firstly, I would like to extend my sincere thanks to my advisor and my co-advisor, who through their support and encouragement, have guided me throughout this process, to bring this thesis to perfection. My advisor Assistant Professor Raif Serkan Albayrak, has always been there with devotion to pose the fiercest and most challenging questions pushing the limits of critical thinking. I am thankful for the enthusiasm and dedication we shared in creating value and contributing to the advancement of social science. My co-advisor Professor Pınar Evrim Mandacı, has always offered her kind and attentive guidance, be it face-to-face on a hot summer day when the university was almost empty or late at nights going to elaborate lenghts to provide valuable timely feedback in the form of emails. Thank you for being the other utmost critical and disciplined supervisor to this work. The third member of my dissertation committee, who has been with me since the first day of this academic journey. Associate Professor, Çağrı Bulut, I value your insightful feedback and never-ending genuine support to this thesis. Thank you for always believing in me and my work and for being a great mentor. A heartfelt thank you goes out to Professor Hasan Fehmi Baklacı, who through his valuable comments on behavioral finance enriched this work. Thank you for taking the time out of your busy schedule and offering your kind suggestions. Another sincere thank you goes out to Assistant Professor Fatma Dilvin Taşkın, who kindly provided interesting feedback and resourceful comments. Thank you for contributing to the fulfillment of this thesis. Secondly, I would like to express my gratefulness to Associate Professor Hakan Kahyaoğlu, whom I met thanks to my co-advisor, Professor Pınar Evrim Mandacı. Thank you for your valuable contributions to important aspect of the econometric modelling framework of this thesis. Thirdly, I would like to thank Barış Turan, my partner in life, for his patience and continuous support. Lastly, I am extremely grateful and indebted to my family, my mother Handan, my father Aydil and my dearest brother Serhan Son. Without their love, support and inspiration throughout my life and during this tiring and challenging process, I could not have accomplished anything I did. This thesis is dedicated to my mother, my role model in life, my best friend and ultimate teacher, the most selfless and perfect human being I know of. She is my rock in my life and the reason that I am what and who I am.

YEMİN METNİ

Doktora Tezi olarak sunduğum "The Effect of Internet Search Volume (ISV) on Stock Return Volatility" adlı çalışmanın, tarafımdan bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım eserlerin bibliyografyada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

..../..../.......

Semen SON

İmza

ÖZET

Doktora Tezi

İNTERNET ARAMA HACMİNİN HİSSE SENEDİ GETİRİSİ OYNAKLIĞI ÜZERİNDEKİ ETKİSİ

Semen SON

Yaşar Üniversitesi
Sosyal Bilimler Enstitüsü
İşletme Doktora Programı

Bu çalışmanın temel amacı, internet arama hacminin, bireysel yatırımcı hissiyatını temsil eden bir dışsal değişken olarak kullanılması suretiyle, hisse senedi getirisi oynaklığı üzerindeki etkisini araştırmaktadır. Çalışmada internet arama hacmi değişkeni ile birlikte işlem hacmi değişkeni de ele alınarak her ikisinin oynaklık üzerindeki etkisi incelenmiştir. Çalışma, NASDAQ, NYSE and BIST piyasalarında işlem gören 52 ABD ve Türk şirketini ele almakta ve Ocak 2004-Eylül 2013 dönemini kapsamaktadır. ABD şirketleri ayrıca e-ticaret şirketi olan ve olmayan şirketler olmak üzere iki alt gruba ayrılmıştır. E-ticaret şirketleri ayrı bir örneklem olarak incelenmiştir. Bu şekilde yatırımcı ve müşteri ayırt edilerek isim-bazlı internet arama sorgularının bireysel yatırımcıya ait olduğu gösterilmiştir. İki farklı koşullu ortalama tanımıyla oluşturulan GARCH(1,1) modellerinin sonuçları bu prosedürün uygunluğunu teyit etmektedir. Internet arama hacmi ile hisse senedi getiri oynaklığı arasındaki ilişkinin yönünü belirlemek amacıyla Granger nedensellik testi kullanılmıştır. Tüm testler şirket bazında yapılmıştır.

Ampirik bulgular, çalışmada bireysel yatırımcı hissiyatını ölçmede kullanılan işlem hacmi ve internet arama hacminin tek tek etkisini her iki ülke piyasası hisse senetleri için de ortaya koymaktadır. Bunlardan geleneksel bir hissiyat değişkeni olan işlem hacmi değişkeninin etkileri ABD senetlerinde daha belirgin olarak gözlemlenmektedir. Bunun yanısıra, çalışmada işlem hacmi ile internet arama hacminin ortak etkisinin varlığı da gözlemlenmektedir. İki değişken, sırayla, şartlı oynaklık denklemine dahil edildiğinde, oynaklık süreğenliğinin azaldığını ancak hisselerin hata terimlerinin büyük çoğunluğunda GARCH etkisinin devam ettiği ortaya konulmaktadır. Granger nedensellik testleri bağlamında zamansal bir sıralamadan söz etmek mümkün değildir.

Bu çalışma, diğer çalışmalardan farklı olarak yatırımcı hissiyatını ölçmede sıkça kullanılan işlem hacmi değişkeninin yanında literatürde çok sık rastlamadığımız internet arama hacmi değişkeninin hisse senedi oynaklığına etkisini incelemektedir. Gözlemleyebildiğimiz kadarıyla internet arama hacminin Türk hisse senedi piyasasının oynaklılığına etkisini inceleyen bir çalışma bulunmamaktadır. Böylelikle çalışmadan elde edilen sonuçlar gerek ulusal gerekse uluslararası davranışsal finans literatürüne önemli katkılar sağlamaktadır.

Anahtar kelimeler: GARCH, Koşullu Oynaklık, Granger Nedensellik, İnternet Arama Hacmi

ABSTRACT

PhD Thesis

THE EFFECT OF INTERNET SEARCH VOLUME (ISV) ON STOCK RETURN VOLATILITY

Semen SON

Yaşar University
Graduate School of Social Sciences
PhD in Business Administration

The aim of this study is to investigate the impact of investor sentiment on stock return volatility by using a novel proxy as exogenous variable: internet search volume. Internet search volume is also analyzed together with trading volume and the effects of both on stock return volatility are being reported. The data set consists of a total of 52 U.S. and Turkish companies belonging to NASDAQ, NYSE and BIST markets for the period of January 2004-September 2013. U.S. companies are divided into two groups of e-businesses and non-e-businesses. This procedure justifies the attribution of name-based search queries to individual investor sentiment by differentiating between customers and investors. GARCH(1,1) model results obtained from two different conditional mean specifications confirm the correctness of this reasoning. All models along with Granger causality testing are applied on a company basis.

Empirical findings show significant investor sentiment effects for each security listed on both countries' exchanges. Also, evidence on the significant effect of internet search volume with a second, traditional investor sentiment measure, trading volume, is presented while the latter is relatively more pronounced in securities belonging to U.S. markets. As the model becomes more nested, there is an evident decline in volatility persistence, however in no case, on the average, the GARCH effect vanishes. No generalizable Granger causal relation between these two exogenous variables is uncovered. Overall, findings suggest that stock return volatility is not only conditional upon its previous values but also on investor sentiment and information flow.

The study differentiates itself from others in that it uses trading volume together with a novel investor sentiment variable, ISV. To the best of our knowledge, there exists no Turkish study that studies the effects of ISV on stock return volatility. The results contribute to both, global and local behavioral finance literature.

Keywords: GARCH, Conditional Variance, Granger Causality, Internet Search Volume

THE EFFECT OF INTERNET SEARCH VOLUME (ISV) ON STOCK RETURN VOLATILITY

TABLE OF CONTENTS

CHAPTER 3
METHODOLOGY

CHAPTER 4
THE MODEL, DATA AND EMPIRICAL FINDINGS

CHAPTER 5
CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF FORMULAS

LIST OF ABBREVIATIONS

ARCH	Autoregressive Conditional Heteroskedasticity
BIST-100	Borsa Istanbul 100 Index (former IMKB- /ISE-100)
CAPM	Capital Asset Pricing Model
EMH	Efficient Market Hypothesis
F/F 3 Model	Fama-French 3 Factor Model
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
ISV	Internet Search Volume
MDH	Mixture of Distributions Hypothesis
NASDAQ	National Association of Securities Dealers Automated Quotations
NYSE	New York Stock Exchange

INTRODUCTION

"Now, the question is no longer, as it was a few decades ago, whether investor sentiment affects stock prices, but rather how to measure investor sentiment and quantify its effects."

Malcolm Baker and Jeffrey Wurgler (2007: 130)

Are asset prices predictable? The search for an answer to this long-debated question has spurred a remarkable outpouring of literature encompassing various disciplines grounded in theories extending over centuries.

It was, especially during the second half of the twentieth century, that researchers ardently conjectured hypotheses based on premises of the rationality of the decision maker. The idea was simple: If investors were utility maximizing rational decision makers and had free access to all information on securities, the market price of a security would equal its expected value. Thus, no investor, using a certain trading strategy, would be able to make superior profits. Consequently, asset prices models of the time, considered only one risk factor to be influencing expected security prices: the risk or the movements of the market, commonly referred to as "systematic risk". Any residual or "idiosyncratic" risk, was presumed to be negligible because rational investors would supposedly diversify it away. This general idea of the informational efficiency of markets and the existence of the rational decision makers having homogeneous expectations regarding stock prices prevailed till the early 1980s.

It was initially through empirical findings suggesting auto-correlation inherent in previous values of a security, that the non-predictability of asset prices came to be challenged. These developments instigated a bulk of research referred to as "market anomalies". Researchers thus, started reconsidering the basic premises of rationality and informationally efficient markets.

Some argued that patterns of predictability may be traced to irrational traders acting in concert and misinterpreting information. Others, believing in perfect markets, while acknowledging the counter-evidence to non-predictability, argued that any irrational movement would be arbitraged away by rational institutional traders thus rendering irrational behavior trivial to the stock price formation process. Yet, another strand of literature, posited that there are limits to arbitrage and that prices may deviate to such extremes that even rational traders may no longer be willing, or have the capacity to, make counter-trades. While irrationality was attributed mostly to individuals, there are also studies who accuse institutions of acting upon “noise”.

In parallel to empirical findings, finance researchers started to become aware and acknowledge studies on investor sentiment from behavioral psychology literature. Therein investors trading on noise rather than information were not necessarily considered as irrational. Rather, the concept of bounded-rationality, exemplifying the limits to human memory and capacity was embraced. In that regard, non-rational investor behavior was attributed to certain ways of behaving or “heuristics”. Next to various other heuristics, investor sentiment emerged as such heuristic-driven explanation in many finance papers viewing “behavioral finance” as a sub-discipline.

Overall, the predominant view of rational investors operating in informationally efficient markets marked by no arbitrage opportunities rendering any strategy geared towards the prediction of stock prices valueless, was replaced with the recognition of limits to arbitrage, investors being rationally-bounded and acting together based on their sentiment.

The econometric modelling literature, too experienced changes over the course of decades. Seminal studies on stock price volatility used to consider the residual or the noise term as displaying a constant variance. Thus the ordinary least squares regression (OLS) was used for volatility modelling purposes. However, once empirical findings demonstrating that stock prices contain autocorrelation started populating literature new models were developed factoring in autoregressive terms.

While believers in informationally efficient markets still exist, there is a growing dominance in literature of believers in the complementary value of behavioral explanations to financial phenomena.

Studies presenting investor sentiment as a variable that needs to be analyzed in that realm, have used various direct and indirect measures such as surveys, firm ratios and trading volume among others. However, with the growth of technology and the availability of internet search queries data offered as a free database, these traditional measures are likely to be complemented by a novel proxy: Internet Search Volume (ISV).

The citation by Baker and Wurgler (2007) underscores that traditional, or classical, theories of stock price formation need to be supplemented by measures of sentiment. This thesis analyzes ISV as proxy for investor sentiment affecting the conditional movements of stock prices, in isolation and combination with one of the most popular traditional proxies, trading volume.

CHAPTER 1
BACKGROUND

1.1. Motivation

Following Keynes' statement in 1937 that human beings are guided by animal spirits, social science researchers have since been wondering about the cognitive processes affecting decision making. Translated into the financial markets setting in particular, the basic premise of the investor who is taking decisions rationally, is challenged by the rationally-bounded investor. The rational investor, having free access to all available related information, is assumed to be able to fully-diversify his portfolio[1]. As such, he would, for instance, prefer to buy index funds instead of wasting energy on stock picking. However, not all investors are rational decision makers, furthermore not all investors have access to all information or are free of financial limitations.

This awareness has led us to think about whether individual investors can influence stock prices. Thus; our starting ground becomes: *Does the individual investor impact movements of stock prices?*

Although data on the percentage of institutional stock ownership is present, no clear-cut line between the rational "institutional" (mutual funds, hedge funds, pension funds, private equity funds) and the "individual" investors can be drawn, since the latter are also invested in those institutions. Therefore it is difficult to determine the percentage of individual investors invested in a certain security.

Thus; since it is impossible to observe the consequences of action of individual investors directly, we are in need of a quantifiable proxy. In that realm ISV data presents itself as a valuable source of information.

In practice, noise is observed in stock prices. Since the early 1980's, literature researching potential causes and modelling stock volatility has grown so much that it

[1] Full diversification is theoretically not possible, for further discussion see Elton and Gruber (1977)

has become a gigantic topic in finance. While early research has concentrated on stock market volatility as a whole, the attention has shifted to encompass the analysis of the volatility of individual stocks.

The motivation of this thesis comes from: (a) empirical evidence that counters the premises of informational efficiency of markets and non-predictability of stock prices (b) theoretical explanations of such based on behavioral heuristics (c) the ongoing quest in academic research that seeks to uncover factors underlying aggregate volatility and noise (d) the bulk of investor sentiment literature along with trading volume and the mixture of distributions hypothesis (e) technological developments offering access to sources of financial time series data, ISV, whose value to volatility studies, remains, as of yet, mostly unexplored.

1.2. Aim and Scope

The aim of this thesis is to uncover the effect of a newly emerging proxy of investor sentiment; ISV, on stock return volatility and analyze its interaction and effect on conditional volatility along with trading volume. In this process, alternative mean and variance specifications are formulated. The reason we use the additional variable, trading volume, is that in the strictest sense, the Efficient Market Hypothesis (EMH) argues that prices are always correct and reflect all available information. Thus, there should be little disagreement as to the price formation. Then, there should be very little trading, based on this reasoning. If there are differences in beliefs regarding information, only then would there be trading beyond liquidity needs. So, EHM does not accept heterogenous beliefs, but behavioral finance, does. Trading volume, is predominantly said to proxy the pace of information flow into the market and some researchers go as far as to argue that it accounts for the volatility clustering effects in underlying securities. So, if we analyze trading volume in conjunction with ISV, our results may shed further light to the volatility literature.

The empirical framework is based on theoretical approaches of behavioral finance focusing on noise trading and investor sentiment.

The data used in the final sample belongs to a broad selection of stocks from the U.S. National Association of Securities Dealers Automated Quotations (NASDAQ) and New York Stock Exchange (NYSE) markets and the BIST-100 Index of Borsa Istanbul. ISV data is obtained from the Google Trends website.

To double-check the adequacy of using company names as keywords in representing ISV data and, secondly, to develop the argument that ISV data used in this study is more representative of the financial investor rather than a potential consumer, a separate sample of e-businesses is used.

Empirical tests involve conditional volatility modelling using GARCH methodology proposed by Bollerslev (1986). The temporal relationship of stock return with ISV and trading volume variables are investigated through Granger Causality analysis (Granger, 1969).

ISV data provided by Google is at a weekly frequency and available only since January 2004. Thus, the time frame of analysis is between January 2004-September 2013 for most stocks, with the exception of a few who have fewer available observations.

1.3. Significance

This thesis belongs to a recently emerging strand of finance literature on the importance of investor sentiment measured through ISV. The usage of ISV information in explaining financial phenomena is based on the seminal study of Da, Engelberg and Gao (2011). To the best of our knowledge, this is the first thesis that uses ISV data in this particular context.

Presenting ISV as a proxy for investor sentiment and determining its effects on stock return volatility along with trading volume effects and in isolation, has an important implication for asset pricing literature: the potential need to integrate an additional explanatory variable.

Furthermore, the inclusion of data belonging to stock markets from two countries; the USA and Turkey, allows for an original comparative study.

Details of the research methodology and findings of this thesis are relevant to all players in financial markets, including regulators and investors, and practitioners in derivatives markets as well as academicians in finance.

1.4. Structure of Thesis

The tesis consists of five chapters. Following the present chapter 1, chapter 2 discusses related empirical and theoretical studies starting with classical finance concepts of the EMH, the Random Walk Hypothesis, Rationality and Expected Utility Theories. The chapter moves through challenges to efficient markets, both, on theoretical grounds embodied in the behavioral finance literature, and, empirical findings like market anomalies. Conditional volatility studies are briefly presented. The chapter concludes with the investor sentiment literature, the mixture of distributions hypothesis-trading volume and ISV literature followed by the research questions. Chapter 3 provides a basic discussion on financial time series data and volatility (conditional volatility) modelling. Chapter 4 presents the empirical model along with various mean and variance specifications derived from a synthesis of the preceding two chapters. Data and sampling procedures are given in detail followed by the results of empirical anaysis. And lastly, chapter 5 discusses the findings, limitations and contributions of the study and provides suggestions for further research.

CHAPTER 2
LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The conceptual framework of related literature goes back to the early 1600s and is grounded in mathematics, probability and philosophy.

This chapter starts with a discussion of EMH and asset pricing models and is followed by an account of the challenges it has encountered. The concept of risk underlying stock returns embodied by the concept of volatility is the focal point of this chapter and is disaggregated between systematic and idiosyncratic components.

As underlying assumptions related to decision making processes of investors change, so does the shift of attention to modelling idiosyncratic risk through conditional volatility models.

The search for the source of idiosyncratic risk and its behavior has led to the rise of noise trader models attributing this volatility component to a concerted effort on part of non-rational investors. In time, as opposed to classical beliefs regarding its randomness, certain patterns were discovered and the quest for uncovering the reasons of such, began.

Subsequently, measures seeking to proxy non-observable investor sentiment have emerged, one of such is the popularly used trading volume proxying information flow. ISV is proposed to be a novel addition. To this day, there are only a few studies on ISV and its effects on stock returns or stock return volatility. All of them are detailed at the end of this chapter.

2.1. Classical Finance

The most important developments contributing to the birth of classical asset pricing models have been Expected Utility Theory, Rational Expectations Theory and the EMH.

The theoretical foundation of asset pricing models in the era of classical finance rests on the assumptions of Expected Utility Theory, which puts risk and return as central issues of the investment decision making process. Being rather prescriptive in nature, Expected Utility Theory is a theory of choice under uncertainty for a single decision-maker, based on strict assumptions about preferences. Its roots can be traced back to the explanations of why believing in God is rational; put forth by Blaise Pascal (1670). Among other important contributors are Bernoulli (1738), Feller (1950) and von Neumann and Morgenstern (1947).

Rationality, on the other hand, being regarded as a rather normative concept, suggests or even dictates certain actions. The Rational Expectations Hypothesis can be traced back to Muth (1961), who proposed that asset prices depend partly on what prospective buyers and sellers expect them to be in the future. Basic premises of the rational expectations paradigm assumes that all investors are identical, utility maximizers and their predictions are accurate. Thus, theoreticians from this school of thought posit that outcomes do not show systematic differences from people's expectations of such.

Expected Utility and Rational expectations are the two main pillars upon which rests the EMH.

2.1.1. The EMH

The idea of the impossibility of predicting asset returns if asset prices incorporate all relevant information goes back to the works of Bachelier (1900), and was formalized by Mandelbrot (1963) and Samuelson (1965), who demonstrated that asset prices in well-functioning markets with investors holding rational expectations should follow a generalized form of a random walk known as a submartingale. Empirical evidence was provided by studies such as Kendall and Hill (1953), Osborne (1959), Alexander (1961), Fama (1963, 1965).

Based on the rational expectations framework, Fama (1970) proposed the concept of information efficiency of stock markets in the context of what kind of information is factored in stock prices. Closely associated with and preceding the

EMH, is the Random Walk Hypothesis popularized by Malkiel (1973), whereby stock prices are posited to follow a random pattern and, thus, are serially uncorrelated rendering any analysis of their past prices valueless.

Fama (1970) distinguished between three forms of EMH based on their informational efficiency: (a) the weak form, (b) the semi-strong form, and (c) the strong form. The semi-strong form of EMH has formed the basis for most empirical research. More recent research has expanded to encompass tests of the weak form as well. If no profit can be made through technical analysis then the weak form holds. If analyzing publicly available information, such as annual reports, does not elicit superior returns then that particular market is said to operate in the semi-strong form. The EMH in its strictest form, states that stock markets are very efficient in incorporating all information (information on past values, stock fundamentals and private information) swiftly. Accordingly, even holders of inside information should not be able to make superior returns. Therefore, in the EMH world, it is impossible for investors to beat the market by analyzing past price movements and stock fundamentals since they are already reflected in the prices.

Empirical studies testing that stock prices follow a random walk have used two tests related to: Firstly, that technical analysis is of no use since stock prices are serially uncorrelated and do not exhibit a repeat pattern. Secondly, that analyzing stock fundamentals will give no information about the intrinsic value of a particular security. Pioneer studies providing empirical evidence for historical independence of stock prices and showing that fundamental analysis was of no value were Fama, Fisher, Jensen and Roll (1969), Jensen (1968), LeRoy (1973).

2.1.2. Asset Pricing Models

Theoretical research on the pricing of securities can be traced back to Markowitz (1952, 1959) where the investor's aim is to maximize expected return at a given level of risk. The Capital Asset Pricing Model (CAPM), being independently developed by Sharpe (1964), Lintner (1965) and Mossin (1966), assume that investors use the Markowitz logic of portfolio formation. As an extention, the CAPM introduces the concept of the risk-free rate and states that expected returns equal the risk free rate plus a linear function of its tendency to covary with the market portfolio. Since investors are assumed to be rational decision makers, the only risk that needs to be considered is the systematic risk associated with the market porfolio, any other residual idiosyncratic risk can be diversified to a minimum level.

According to the static CAPM, the expected return E(R_i) of a given financial asset i is presented as:

$$E(R_{i,t}) = R_f + \beta_i (R_{m,t} - R_f) + \varepsilon_{i,t} \quad (1)$$

Where, R_f is the risk-free rate, E(R_m) is the expected return on the market portfolio (i.e., a portfolio of all assets in the economy), ε_i is the residual term and β_i is the sensitivity to systematic risk, which should be compensated by a higher rate of return, equal to the covariance of asset i with the market portfolio (the "beta" of the stock).

In the mid-1960s, this model provided a good explanation of asset prices. However, these explanations received criticism towards the end of the 1970s, with the applications of tests using time-series regressions of stock returns on index returns to generate estimates of stock-specific betas.

The development of the CAPM, led to the Joint Hypothesis problem; which simply states that findings that forecast errors are possibly predictable does not necessarily mean that markets are inefficient. The asset model itself might have been incorrrectly specified. However, an asset-pricing model cannot be tested easily without making the assumption that prices rationally incorporate all relevant available information and that forecast errors are unpredictable.

Fisher, Jensen and Roll (1969), who also introduced the event studies methodology, tackled the Joint Hypothesis model by using "The Market Model" to capture the variation in expected returns as shown below:

$$E(R_{i,t}) = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t} \quad (2)$$

Here $R_{m,t}$ stands for the current overall market return, and α_i and β_i are estimated coefficients from a regression of realized returns on stock i, $R_{i,t}$, on the overall market returns using data before the event. Assuming that β_i captures differences in expected return across assets, $\varepsilon_{i,t}$ represents the residual idiosyncratic noise.

With the assumption that stock returns should be unpredictable, idiosyncractic noise should be uncorrelated across events. This procedure adresses the joint hypothesis problem and isolates the price development of stock *i* from the impact of general shocks to the market.

2.1.3. Criticism and Challenges to CAPM and EMH

A seminal paper by Roll (1977) criticized tests of the CAPM, demonstrating that any valid CAPM test presupposed complete knowledge of the market portfolio. In CAPM theory, the market portfolio contains every individual asset in the economy, including human capital, and is, thus, unobservable. Using a stock market index as a proxy for the market portfolio, as commonly used by previous tests, would therefore lead to biased and misleading results.

Pioneers in shaking the existence to EMH were Shiller (1981) and LeRoy and Porter (1981). Shiller (1981) was the first to attribute his empirical findings of excess volatility to optimistic or pessimistic market psychology. These studies were followed by Schwert (1989), who suggests that volatility of stocks increase during recessions and attributes this movement to operating leverage. The market crash of 1987 was the turning point and eventual demise of EMH.

Contradictory empirical findings against the general applicability of the EMH (commonly referred to as "market anomalies") have followed. These rest on a series of test investigating whether publicly available information used in fundamental analysis can be used to improve returns. Oft-studied anomalies include the "Small-Firm Effect", where it is shown that small firms tend to earn abnormally high returns over long time periods (Reinganum, 1983). The "January Effect" or Turn-of-the-Year Effect", another prominent research topic suggested by Roll (1988), arguing that at the turn of the year stock prices experience an abnormal and predictable rise in prices that is inconsistent with random-walk behavior. French (1980), analyzing daily returns for stocks in 1953-1977, argues that there is a tendency for returns to be negative on Mondays, an anomaly called the “Monday Effect or “Weekend Effect”. Another anomaly called the “P/E Ratio Effect” is suggested by Basu (1977) arguing that low P/E ratio stocks earned a premium during 1957-1971 and Campbell and Shiller (1988) demonstrate the predictive power of P/E ratios. The Over/Under-Reaction Anomaly puts forth that stock prices overreact to certain events such as current changes in earnings (DeBondt and Thaler, 1985). Other numerous anomalies discovered over the years are the Neglected Firm Effect (Arbel, 1985), the Weather Effect (Hirshleifer and Shumway, 2003), and the Book-to-Market Effect (Fama and French, 1992).

Tests developed to determine cracks in the strong form of the EMH, geared towards detecting whether insiders could make superior profits, was adressed by researchers such as Jaffe (1974) and Givoly and Palmon (1985), but are relatively few in number.

Grossman and Stiglitz (1980) went as far as to argue that the existence of perfectly informationally efficient markets is impossible, since if markets are perfectly efficient, there is no profit to gathering information. Therefore there would be little reason to trade and markets would eventually collapse.

The CAPM has by many researchers been accused of being unrealistic. While there has been research proposing extensions to the basic CAPM like the Consumption-Oriented Intertemporal Asset Pricing Model (Breeden, 1979), other theoreticians like Ross (1976) attempted to remedy the defficiencies of the CAPM with a different model namely the Artibtrage Pricing Theory (APT).

The underlying premise of the APT, being an equilibrium model, is that no arbitrage opportunity should be present in efficient financial markets. As opposed to the single beta CAPM, APT assumes there are *n* number of factors which can cause systematic deviations of returns from their expected values; thus an asset’s expected return is a linear function of its sensitivity to *n* number of common factors. As in the

CAPM, residual risk is assumed to be diversifiable. Both the CAPM and APT are single-period models. Merton (1973), introduced the intertemporal capital asset pricing model (ICAPM) to account for the multi-period nature of financial market equilibrium. The ICAPM is very similar to the APT, with the exception that the first factor is defined explicitly as being related to the market portfolio. A further difference is that ICAPM puts requirements for factors to be included into its equation and thus, has a restrictive nature on the number of factors underlying.

The early excess-volatility findings were also challenged on econometric grounds by Marsh and Merton (1986) and Kleidon (1986), who noted that the test statistics used by Shiller (1979, 1981) are only valid if the time series are stationary. This issue was addressed by Campbell and Shiller (1988) who used the Theory of Cointegrated Processes developed by Granger and Engle (1987), to design new tests of the present-value model that allow the processes generating prices and dividends to be nonstationary.

A relatively newer asset pricing model, integrating all previous findings is the Fama French Three Factor Model (F/F 3) (1993), which is based on the premise that the CAPM beta has practically no additional explanatory power once book-to-market (HML) and size (SMB) have been accounted for. The F/F 3 Model is depicted below:

$$\mathrm{E}(R_{i,t}) = \alpha_i + \beta_{\mathrm{mkt}} R_{m\mathrm{kt},t} + \beta_{\mathrm{HML}} R_{\mathrm{HML},t} + \beta_{\mathrm{SMB}} R_{\mathrm{SMB},t} + \varepsilon_i \qquad (3)$$

The authors, arguing that these two factors capture fundamental risk for which investors demand compensation developed a rational multi-factor interpretation of stock returns. In contrast other researchers construed the significance of these two factors as the effect of market mispricing and investor irrationality following Shiller, Fisher and Friedman (1984). Lakonishok, Shleifer and Vishny (1994) attributed excess return to high book-to-market stocks to underpricing by investors and low book-to-market stocks to investors' overpricing so that they subsequently underperform the market.

A parallel development was the challenge to informational efficiency and the F/F 3 by Jegadeesh and Titman (1993), who discovered the “momentum” in stock prices identifying consistently winning (losing) stocks over a 3-12 months horizon.

A common characteristics of most of these models is the underlying assumption that the investor is not only rational but also processes information efficiently.

Other attacks to the F/F 3, came from Black (1993) who suggested their findings were attributable to data mining, thus, their results would disappear if one were to use another data set over a different time period. Similarly, Kothari, Shanken and Sloan (1995) suggested that the Compustat data used in F/F 3 testing might suffer from survivorship bias and that beta calculations are very sensitive to the frequency of data.

Extensions to F/F 3 such as the Carhart Four Factor Model (Carhart, 1997) includes a fourth factor namely momentum besides the factors including beta, size and book-to-market ratio.

2.2. Behavioral Finance and Prospect Theory

Behavioral finance is the outcome of a broad compilation of works in finance and psychology.

Its roots can be witnessed in the works of Keynes and his concept of “animal spirits (1937), Simon (1955) and March and Simon (1958) putting forth the bounded-rationality principle; where decision making is limited, followed by the Theory of Cognitive Dissonance (Festinger, Riecken and Schachter, 1956), Samuelson’s fallacy of large numbers (1963) and advanced through the introduction of such concepts as “the availability heuristic”, “representativeness, availability and anchoring and adjustment”, “loss aversion”, “framing”, “under/over-reaction”, “herd behavior”, “overconfidence” by Kahneman and Tversky (1973, 1974, 1979, 1981), Daniel, Hirshleifer and Subrahmanyam (1998), Shiller (2000) and Shefrin (2000), respectively. However, it was Thaler's work in 1980 that promoted prospect theory to be used as basis for an alternative descriptive theory in economics.

The main argument of behavioral finance is that deviations of asset prices from their fundamental values not considered by the EMH, can be interpreted as being caused by the presence of investors in financial markets that are not fully rational. Shleifer and Summers (1990) posit that behavioral finance rests upon two pillars: limits to arbitrage, seeking to explain the existence of arbitrage opportunities and investor psychology not necessarily based on fully rational models.

Examples of studies using behavioral explanations for financial phenomena include: Lakonishok, Shleifer and Vishny (1994) who claim that value strategies yield higher returns because these strategies exploit the suboptimal behaviour of the typical investor. Benartzi and Thaler (1995) address the equity premium puzzle of Mehra and Prescott (1985), which refers to the fact that stocks have outperformed bonds by a far greater degree than warranted under the standard expected utility maximizing paradigm. The authors attribute this anomaly to a concept they call "myopic loss aversion"; loss aversion combined with a prudent tendency to frequently monitor one's wealth. On the other hand, analyzing mutual funds behavior, Grinblatt, Titman and Wermers (1997), demonstrated evidence of momentum strategies and herding. A widely recognized heuristic is that of representativeness based on DeBondt and Thaler's (1985) study. They argue that investors become extremely pessimistic (optimistic) about past losers (winners) and, consequently, overreact to both bad and good news leading past loser (winners) to become under (over)priced. Shefrin (2000) following Odean (1998), presents overconfidence as a reason for investors' excessive trading.

One of the major opponents to behavioral finance, is the founder of EMH, Eugene Fama (1998). While admitting that there may be some anomalies that EMH cannot adress, he argues that it should not be fully dismissed and replaced by behavioral finance. Shiller (2003) defends behaviorism against these criticisms especially stressing excess volatility in stock returns and the fact that this phenomena is of yet not been refuted. Thaler (1999) argues that behavioral finance cannot be dismissed due to the theoretical possibility of cognitive biases as exerting influence on asset prices.

The following sections give a more detailed look at potential pillars of investor behavior, mentioned in various finance studies: Bounded-Rationality (Satisficing), Heuristics (Sentiment), and Noise Trading.

2.2.1. Pillars of Investor Behavior

Grether (1992) defines heuristics as as a rule of thumb or decision aid by which individuals may judge likelihood. Shefrin (1999) argues that it is heuristics and frame dependence that lead to market inefficiencies causing prices to deviate from their fundamentals.

Behavioral heuristics assist behavioral finance researchers in explaining issues like why investors fail to diversify, why they sell winners and keep losers, why they trade actively and over/underreact to news. As opposed to the rational expectations paradigm investors are not assumed to be idential and tend not to follow Bayesian rules to form new beliefs as information becomes available. Thus, not all investor predictions are accurate. In practice, investors may be rationally-bounded and display *"satisficing"* behavior (Simon, 1956) due to information limitations. Satisficing is a cognitive heuristic which states that since human beings lack cognitive resources to evaluate all outcomes with sufficient precision and do not know relevant probabilities due to unlimited memory, they cannot optimize. Shefrin (1999) defines a frame as a description. Frame dependence implies that people make decisions that are influenced by the manner in which the information is presented.

Kahneman and Tversky (1972) and Tversky and Kahneman (1971, 1974) introduced the "*representativeness*" heuristic explaining that people make probabilistic judgments based on similarity. Representativeness causes people to give too much weight to recent evidence and too little weight to the base rate or prior odds. It may also lead to extreme forecasts relative to the predictive value of the information available. DeBondt and Thaler (1985) applied overconfidence to their behavioral explanation regarding market pricing. They argued that investors react to both, good and bad news. Thus, overreaction causes past loser stocks to become underpriced and past winners to become overpriced.

Shefrin and Statman (1985) apply Kahneman and Tversky's notion of "*framing*" to the realization of losses and called that phenomenon the "Disposition Effect". Presumably investors are predisposed to holding losers too long and selling winners too early.

Shefrin (1999) suggests that these two studies opened two different avenues for investigating the implications of behavioral phenomena, with one stream looking into security prices and the other into the behavior of individuals.

Kahneman and Tversky in their seminal and remarkable paper on Prospect Theory (1979), question the validity of expected utility theory. Their argument is that when faced with the complex task of assigning probabilities to uncertain outcomes, individuals often revert to the use of behavioral heuristics which in turn lead to systematic biases. Prospect Theory was developed as a more accurate alternative psychological model for decision making under risk, compared to Expected Utility Theory, the latter resting upon the "reality axioms" of von Neuman and Morgenstern (1947).

Prospect Theory replaces the probabilities put forth by Expected Utility Theory by decision weights which assign value to gains and losses (changes in wealth or welfare) rather than absolute magnitudes. In this sense it is rather descriptive as it observes the behavior of individuals rather than normatively dictating what investors *should* do given certain assumptions about them. Accordingly value should be dependent upon two arguments: the asset position that serves as a reference point, and the magnitude of change from this respective reference point.

> *"Many sensory and perceptual dimensions share the property that the psychological response is a concave function of the magnitude of physical change."* (Kahneman and Tversky, 1979: 278)

The above quote describes an essential feature of Prospect Theory that the Value Function proposes. Individuals cannot discriminate a temperature change from 13 degrees to 16 degrees as easily as they can when the temperature changes from 3 degrees to 6 degrees. This recognition is applied to economics and argued that the value function for wealth is concave above the reference point and gets convex below it. Thus, marginal value attained from gains and losses tend to generally decrease in their magnitude. The idea of loss aversion is an outcome of the value function being steeper for losses than for gains.

Following Prospect Theory numerous other behavioral heuristics were applied to finance and popularized through several models. One of them is what Barber and Odean (2008) refer to as the "*Investor Sentiment Model*" in which investors over/underreact to information due to the "overconfidence". Yet another one is the "*Noise Trader Model*" by Delong et al. (1990) which involves investors reacting to irrelevant information, that is they may interpret signals as information whereas they are merely noise.

While there is no common agreement on what constitutes investor sentiment; be it emotions, heuristics or the propensitity to trade on noise, there seems to be an inclination to attribute these characteristics to the individual investor who does not based his decisions on preferences but rather beliefs.

2.2.2. Heuristics

This section provides an overview of a few relevant heuristics to this thesis, the large universe of heuristics is not discussed, for brevity purposes.

(i) Overconfidence: Barberis, Shleifer and Vishny (1998) were the first researchers to model this heuristc under the so-called "Investor Sentiment Model", to explain how investors form their beliefs that are translated as investor under and over-reaction in stock markets. Accordingly, overreaction means that the average return following a series of announcements of good news is lower than the average return following a series of bad news announcements. After a series of good news, investors are observed to become overly optimistic that the future news announcements will be also good and, hence, overreact causing an overvaluation of the firm's stock price. However, they will be experience anxiety once subsequent news announcements do not confirm their prior optimism, which leads to a price decrease below its fundamental value and, then, to poor returns. This reversion means that the overweighting of bad (good) information leading to a decrease (increase) in prices below (above) their fundamental value is corrected in the subsequent period.

Among other researchers, who have attributed excess trade to overconfidence are Odean (1998), Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001),

(ii) Confirmation Bias: Confirmation bias is defined by Nickerson (1998) as "the seeking or interpreting of evidence in ways that are partial to existing beliefs". This heuristc implies that an investor would be more inclined to search for information that supports his or her original idea about an investment rather than seek out information that contradicts it. Consequently, this bias can often result in faulty decision making. Pouget and Villeneuve (2012) study confirmation bias and propose a dynamic model where some investors are prone to it. In a model with public information only, this assumption is said to provide a rationale for the volume-based price momentum documented by Lee and Swaminathan (2000). A study by Park et. al. (2010) conjecture that investors would use message boards to seek information that confirms their prior beliefs making them overconfident and adversely affect their investment performance. Their analysis of 502 investor responses from the largest message board operator in South Korea supports their hypothesis that investors exhibit confirmation bias when they process information from message boards.

(iii) Framing: Shefrin (2000) suggest the term frame dependence means that the way people behave depends on the way that their decision problems are framed. In classical finance, framing is argued to be transparent such that investors can see through all the different ways cash flows might be described. As a bias, it is based on cognitive and emotional factors which influence the mental organization process of information and how the outcomes are coded into gains and losses.

(iv) Herd Instinct (Herding): There is a natural instinct for people to become part of a group so they tend to herd together. Moving with the herd magnified psychological bias inducing the person to decide based on his feelings instead of performing independent analysis. In behavioral finance, herding presents another strand of literature looking to explain price deviations from their fundamentals. For instance, Barber, Odean, Zhu (2009) findings on U.S. stock markets over the period 1983 to 2001 suggest that individual investors predominantly buy (sell) the same stocks as each other contemporaneously. Redding (1996) gives a detailed overview of noise trading and the herding literature.

(v) Sentiment: Is the reflection of heuristic-driven bias. Brown (1999), describing the relation between noise and investor sentiment, argues that if noise traders affect prices, the noisy signal is sentiment, and the risk they cause is volatility then sentiment should be correlated with volatility. Studies such as Kumar and Lee (2006) provide powerful and consistent empirical evidence that stock prices are affected by sentiment risk in contrast to studies such as Sias, Starks and Tinic (2001) who argue that financial markets do not price cognitive factors. Beer, Watfa and Zouaoui (2011) investigate whether noise trader risk (also called “sentiment risk”) is valued by the stock market. The authors find that the impact of sentiment risk on stock returns is more associated with certain types of stock like small stocks, growth stocks, young stocks, unprofitable stocks, lower dividend-paying stocks, intangible stocks and high volatility stocks.

2.3. Noise Trading Models

Sentiment is what leads investors not be necessarily irrational, but to form systematic biases in the way they believe and causes them to trade on non-fundamental information. The relation between sentiment and asset pricing has become popular starting with Black (1986) followed by various other studies like Delong et al. (1987), Barberis, Shleifer and Vishny (1998), Daniel, Hirshleifer and Subrahmanyam (2001). All of these studies assume the categorization of investors in generally two groups: the informed traders acting rationally on information and uninformed noise traders relying on their “irrational” sentiment. However, a commonly agreed upon clear-cut categorization and the source of noise trader risk remains unanswered. For instance, Baker and Wurgler (2006) put forth that mispricing based on sentiment is caused by a combination of uninformed demand of some investors, the noise traders, and a limit to arbitrage.

The notable Noise Trader Model by Delong et al. (1990) is based on the assumption that there are noise traders and sophisticated investors. The model is describing the impact of noise trading on equilibrium prices. The price deviations from fundamental values are a result of unpredictable investor sentiment. Arbitrageurs who trade against mispricings are faced with the risk of investor sentiment becoming more extreme and thus moving prices even further from their fundamental values. Eventually, arbitrageurs fail to entirely eliminate these sentiment-related mispricings, leaving noise trader risk factored into the price formation process. Instead of expanding upon what the source of the false belief or sentiment about irrelevant information they base their trading decisions on the implications of the existence of noise traders. They point out that studying irrational behavior does not always require specific content and that by mere observation of the effect of the unpredictability of irrational behavior on the opportunities of rational investors, something can be learned.

Models of investor sentiment base their assumptions on various heuristics. Barberis, Shleifer and Vishny (1998) base their model on representativeness and conservatism as opposed to Daniel, Hirshleifer, and Subrahmanyam (1998) whose model attempts to reconcile the empirical findings of overreaction and underreaction basing them on heuristics such as overconfidence and self-attribution.

Over the course of three decades, the effect of investor sentiment on asset returns has been subject to detailed empirical testing. Many studies (such as Lee, Shleifer, and Thaler, 1991; Lee et al., 2002; Brown and Cliff, 2005; Baker and Wurgler, 2007; Ho and Hung, 2009; Baker, Wurgler, and Yuan, 2012) demonstrate that there is a contemporaneous relationship of such. Apart from individual assets, the effect of investor sentiment on stock market volatility has been explored and evidenced in studies such as Brown (1999).

Lee, Jiang and Indro (2002) describe that underlying noise trader models in finance is the premise that subsets of agents trade in response to extraneous variables that convey no information about fundamentals.

Malkiel (2003) cites other noise models in addition to Delong et al. (1990), like the studies by Campbell and Kyle (1993), Campbell, Grossman, and Wang (1993), or Llorente et al. (2002) that predict that noise trading adds to idiosyncratic volatility above and beyond cash flow news and concludes that retail trading may positively affect volatility if individual investors behave as “noise traders” or “liquidity traders.

While there seems to be a common agreement in literature that investor behavior does have an effect on financial markets, the extent to which is still questionnable. Furthermore, distinguishing among types of investors (institutional or individual) and their relative contributions to EMH-contrarian outcomes is very difficult. Thus, behavioral finance may be viewed as offering complementary arguments in explaining asset price behavior next to traditional models.

2.4. Volatility

The fluctuation or “variance” of stock price returns over a certain time period, is called stock return volatility or simply “volatility”. It is of particular interest to investors, dealers, brokers, regulatory agents, risk managers and company owners. Volatility is often times equated with risk and consequently the more stable stock price returns are, the less riskier they are perceived to be.

The empirical papers of Shiller (1979) and LeRoy and Porter (1981) demonstrating that stock prices and long-term interest rates fluctuate more than predicted by traditional asset models, were followed by Shiller (1981), and Amsler (1984) reporting similar findings of excess volatility.

In general, the return of an asset can be decomposed into its systematic component its idiosyncratic component. Its corresponding volatility can also be divided into two parts:

$$\text{Volatility} = \text{systematic volatility} + \text{idiosyncratic volatility} \quad (4)$$

Xu and Malkiel (2003) assert that stock market volatility as a whole has received considerable press attention during the late 1990s. They further argue that this attention has been misplaced referring to the findings of Schwert (1989), who demonstrated that no long-term upward trends was found in the volatility of the stock market as a whole. The authors draw attention to the fact that volatilities of individual stocks can increase even though the volatility of the market remains stable provided that correlations among stocks are declining. This argument was researched by Campbell et al. (2001), who showed that volatilities of individual stocks had indeed increased during that the period between the 1980s and 1990s, as a result of an increase in their respective idiosyncratic volatilities.

As Xu and Malkiel (2003) put forth, it is very difficult to measure idiosyncratic volatility since it is inherently unobservable and model-dependent.

While the CAPM views idiosyncratic volatility as irrelevant, the challenges posed to traditional asset pricing theories has led researchers to reconsider the role of idiosyncratic risk. Ruan, Sun, Xu (2010) argue that viewed from a theoretical perspective, idiosyncratic risk might be important allowing for some degree of market imperfections. Merton (1987) suggests that idiosyncratic risk could be priced in case investors cannot hold all available stocks.

Goyal and Santa-Clara (2003), building on previous literature, look at average stock risk in addition to market risk and find a significant positive relation between average stock variance and the return on the market. The authors claim that since average stock risk is mostly driven by idiosyncratic risk.

Ruan, Sun and Xu (2010) citing Merton (1987) and Malkiel and Xu (2002), argue that it is not the total of idiosyncratic volatility that needs to be priced but only a portion of such that cannot be diversified away. The authors posit that in practice, about one third of the idiosyncratic risk can be diversified away even if one were to hold a portfolio of two stocks on average. Hence, a proxy to estimate this unobservable portion of volatility should be used. Ruan, Sun and Xu (2010) refer to the priced idiosyncratic component as “signal” and the unpriced part of idiosyncratic volatility as “noise”. However the term “noise” as originally coined by Black (1986), is commonly used to denote aggregate idiosyncratic volatility. Thus, throughout this study idiosyncratic volatility is meant to refer to noise in the aggregate sense of its meaning.

2.4.1. Studies on Volatility Measurement with Exogenous Variables

Initial volatility studies in finance and econometrics have come a long way since the 1950s, when Harry Markowitz used standart deviation as a general measure to demonstrate risk reduction through the benefit of diversification. Bollerslev and Wooldridge (1992), referring to risk as “uncertainty”, explains that although the uncertainty of speculative prices was recognized in literature since Mandelbrot (1963) and Fama (1965), it was with the introduction of the Autoregressive Conditional Heteroskedasticity Model (ARCH) of Engle (1982), when researchers started realizing that volatility in high frequency time series data, such as asset returns, is time-varying.[2]

The ground-breaking ARCH Model has changed the landscape of volatility studies and has received numerous extensions. One of the most important ARCH-type models includes the linear Generalized Conditional Heteroskedasticity (GARCH) Model introduced by Bollerslev (1986).

Explained in statistical terms, until ARCH-family type models became popular, standart regression models, as the univariate equation presented below, assumed the error term “ε”, representing idiosyncratic volatility, to exhibit a constant variance.

$$y = \alpha + \beta x + \varepsilon \tag{5}$$

[2] Besides historical volatility (extracted from return series), there’s also the concept of implied volatility (extracted from option prices) and realized volatility (the sum of squared returns). For a comparative discussion see Koopman, Jungbacker, & Hol (2005).

Where, the expected value *y* is dependent upon a constant α plus a proportion β of the variable *x*. This proportional dependency is subject to an error term ε because it cannot be estimated exactly. It is this error term that has been subject to prominent debates to its calculations. In other words, the variance is considered equal to the expected error squared as denoted by $\sigma^2 = E(\varepsilon^2)$. This assumption, that the expected error does not depend on the size of variable x and is constant, is called homeoskedasticity.

However, once non-constant variance (heteroskedasticity) is established, random walk models dependent upon the aforementioned rational expectations hypothesis, can no longer capture the temporal variation in conditional means (and are also no longer valid to behavioral finance researchers theoretically). Thus the newer ARCH-type models became very popular due to their success in modelling the varying variances that are *conditional* upon their past values.

Engle (1982) was the first researcher to propose a model to capture time varying conditional variance in the error terms instead of ignoring it. ARCH (Autoregressive Conditional Heteroskedasticity) processes remain popular till present day.

In order to accurately describe conditional variance, the need for higher order lags of autoregressive compents paved the way to the construction of the more parsimonious GARCH models by Bollerslev (1986). Therein, the conditional variance is thought be a linear process, consisting of a function of past squared residuals and previous conditional variances. GARCH models are thought of as being symmetric models in that they do not differentiate responses to positive and negative news.

There are many proponents of using GARCH models when modelling financial time series data. Among such, for instance Aybar and Yavan (1998), conclude that asymmetry is not a universal phenomenon and suggest symmetric GARCH(1,1) to be a better fit.

Volatility models related to equity stocks, are often estimated without exogenous variables (Antweiler and Frank, 2004: 1260). However, there are also studies including one or more exogenous variables such as; employment rates, CPI, home sales (Flannery and Protopapadakis, 2002), T-bill rates (Engle and Patton, 2001), money growth (Geske and Roll, 1983), gold prices and discount rates (Mangani, 2009), and exchange rates (Araghi and Pak, 2013). Brooks, Faff and Fry (2001) and Depken (2001) use financial variables, such as firm size and trading volume.

In fact, the relation between the volatility of stock returns and changes in trading volumes measuring the rate of information arrival goes as far back as the early 1970s. A famous model is that by Tauchen and Pitts (1983) combining important features of Clark (1973) and Epps and Epps (1976) which are by now the most-often mentioned studies on the mixture of distributions hypothesis.

Lamoureux and Lastrapes (1990), examing the U.S. stock market, use the GARCH(1,1) model and base their work on the mixture of distributions hypothesis to show that varying flows of information causes the variance of stock prices to vary. Their results mainly indicate that when trading volume is positively related to stock return volatility, ARCH effects tend to decrease and generally become insignificant suggesting that when the rate of information flow is accounted for lagged squared residuals do no longer contain any, or any major, additional information about stock return variances.

2.4.2. Sources of Idiosyncratic Volatility

Under-diversification is commonly agreed upon as being the major cause for idiosyncratic risk, however, while there are rational explanations to the reasons for under-diversification (like taxes, transaction costs, limited resources) there is also the irrationality explanation. For instance Shiller (2000) attributed the internet boom to a concept he called “irrational exubarance”.

Berrada and Hugonnier (2012) explain the deviation of expected returns from those predicted by standart asset pricing models, which they refer to as “the product of the stock's idiosyncratic volatility and the investors' aggregated forecast errors", to investors' lack of complete information.

The internet boom (or the dot-com bubble), where prices of especially high-tech stocks boomed and later on busted in the period covering 1997-2000 is one of the major market anomalies and, thus, has received tremendous academic attention.

For instance, Brandt et al. (2010) and Foucault, Sraer and Thesmar (2011), demonstrate that volatility during the internet boom is positively correlated with the trading activity of individual traders behaving as noise traders while Fink et al. (2005) attribute idiosyncratic risk to investor sentiment and firm age. Corporate variables, are often times cited as determinants of cash flow variability such as in the papers of Comin and Philippon (2006) and Wei and Zhang (2006).

Bennett, Sias, and Starks (2003) ascribe the rise in idiosyncratic volatility to an increase in institutional ownership, and especially the increased preferences of institutions for small stocks. Chichernea, Petkevich and Reca (2012), posit that short-term (long-term) institutional ownership is positively (negatively) linked to idiosyncratic volatility. In the empirical study by Greenwood and Thesmar (2011) stocks with very high percentages held by a few institutions are found to exhibit high volatility. Institutional ownership, as in the studies of Bohl, Brzeszczynski, Wilfling (2009) and Cohen, Gompers and Vuolteenaho (2002) is shown to have a stabilizing effect on stock prices. Anton and Polk (2010) explain the cross-sectional variation in return correlation through the degree of common institutional ownership.

Overall, literature has attributed many reasons for idiosyncratic volatility. Next to market specific reasons such as listing requirements, corporate variables such as changes in company policy, the noise trader approach is mentioned in the bulk of volatility studies. However, who these noise traders are, and in particular if they are institutional investors or individuals, and the specific reasons behind the noise they create is not universally agreed upon.

Lee, Shleifer, and Thaler (1991), assume that noise traders are identifiable with individual investors. Barber and Odean (2008) demonstrate evidence that it is the individual investors that display attention-based buying behavior on certain days.

Shleifer and Vishny (1997) divide the financial market up among three types of investors: noise traders, arbitrageurs and investors in arbitrage funds who do not trade on their own. In their model both arbitrageurs and their investors are fully rational. Risk-neutral arbitrageurs take positions against the mispricing generated by the noise traders. In reality, arbitrage resources are heavily concentrated in the hands of a few investors that are highly specialized in trading a few assets, and are far from diversified. As a result, these investors care about total risk, and not just systematic risk.

Kaniel, Saar and Titman (2005) argue that institutions differ from individual with respect to their size and sophistication, however Schmeling (2009) posits that there is considerable differene as to how these two groups differ and which effect they exercise on price formation and market liquidity. On the other hand, bulk of papers inspired by Black (1976) exist that views individiual investors as irrational noise traders. In those contexts informed and irrationality terms are synonomous with being "smart and sophisticated" and "dumb and unsophisticated", respectively.

Added to these debates there is also empirical evidence put forth by Sias (2004) that institutions deliberately herd in and out of stocks and rely on momentum-style trading strategies.

Lease, Lewellen, and Schlarbaum (1974) argue that there is extensive evidence showing that investors do not hold the market portfolio and prefer to hold a single stock or a portfolio of a few stocks. On the other hand, investors that do diversify, according to Jensen (1968), prefer high fee-charging stock-picking mutual funds who nevertheless, are not able to beat the market.

There are also studies that do not attribute importance to them and posit that they can be ignored while explaining the process of price formation. The latter strand of studies as explained by Delong et al. (1987), is based upon Friedman (1953) and Fama (1965).

Fama (1965) argues that irrational noise trading is countered by rational arbitrageurs trading against them and consequently drive prices close to fundamental values and eventually noise traders losing money to sophisticated aribtrageurs would, according to Friedman (1953), be driven out of the market. In contrast Black (1986), who does not spell out clearly who these noise traders are argues that “if there is no noise trading, there will be very little trading in individual assets” (Black, 1986: 529-530). Kyle (1985) associates noise traders with random aggregate demand and no persistent or predictable influence on stock prices.

Shefrin and Statman (1985) divide noise traders into two categories: traders who commit cognitive errors, while information traders are free of cognitive errors.

Bloomfield, O’Hara and Saar (2009) attempt to provide an explanation for the confusion that exists in market-microstructure versus the limits-to-arbitrage literature about the different interpretations of the term “noise traders”. Accordingly, researchers in the former strand like Kyle (1985), use the terms “noise traders” and “liquidity traders” interchangeably to describe traders who do not possess fundamental information and trade for hedging or liquidity purposes. The limits-to-arbitrage supporters, in contrast, attribute noise trading activity to reasons other than fundamental information, hedging or liquidity shocks. This is said to be exemplified in the work of Shleifer and Summers (1990), who explain that the term “noise traders” refers to behavioral causes not taken into account of by standard explanations, and forms the bases of their “noise trader approach to finance”.

The existence of various descriptional studies show that there is no clear cut line between informed traders and noise traders. Therefore the premise that there is, idiosyncratic risk in stock returns, should serve as a commonly agreed upon starting point.

2.4.3. Measuring Idiosyncratic Volatility

By now literature on idiosyncratic volatility and its relation especially to total variance, has accumulated, and it is apparent that various methods for defining this volatility component are being preferred. Since idiosyncratic volatility is the residual term of a mean equation, the specification of the latter is very important. Table 1 presents selected studies which define idiosyncratic volatility through various approaches and reports their findings in relation to expected returns.

Table 1
Empirical Evidence on Idiosyncratic Risk and Return

Study	Sample Period	Idiosyncratic Risk Definition	Expected Volatility Measure	Findings
Lintner (1965)	1954-1963	CAPM residuals	Lagged	PR
Lehman (1990)	1931-1983	CAPM residuals	Lagged	PR
Ang et. al (2006)	1963-2000	F/F 3 Residuals	Lagged	NR
Eiling (2006)	1959-2005	CAPM Residuals	EGARCH	PR
Huang et al. (2007)	1963-2004	F/F 3 Residuals	EGARCH	PR
Brockman & Schutte (2007)	1980-2007	F/F 3 Residuals	EGARCH	PR
Bali & Cakici (2008)	1963-2004	F/F 3 Residuals	Lagged	-
Fu (2009)	1963-2006	F/F 3 Residuals	EGARCH	PR

Adopted from Kotiaho (2010). PR indicates positive, NR indicates negative and "-" indicates no relation.

There are numerous ways to measure idiosyncratic volatility, a common approach is to construct estimates by direct adjustment of total risk for the variation in non-diversiable risk factors identified through models such as CAPM or F/F 3. However, there is still no commonly agreed upon approach. For instance, Fink et al. (2005), argue that since they are focusing on the time-series properties of large portfolios of firms, it is unlikely that methodological differences will affect their main results.

2.5. Proxies of Investor Sentiment and Trading Volume

A prominent study by Baker and Wurgler (2007), explains that a proxy is an observable phenomenon that serves as an exogenous shock in investor sentiment leading to a chain of events affecting patterns in security pricing. It may be the case that it first manifests itself in investor beliefs that could be surveyed, which later on translate into observable trading patterns. Moreover, limited arbitrage causes these demand pressures to lead to mispricings, which in turn may be picked up through benchmarks of fundamental value such as the book-to-market ratio. Mispricing may induce informed responses by insider, like corporate executives holding inside information and having the ability to act on it to influence the leverage situation of the firm.

However, as Baker and Wurgler (2007) point out, this chain is prone to confounding influences like surveys not being an exact illustration of how people do behave versus how they respond to such. The difficulty of using trades, on the other hand, is that they net to zero (there’s a buyer and a seller to each trade), thus, using this measure taking a stand on the identity of irrational investors. And corporate executives may like to change the debt to equity structure of their firms for many reasons other than inside information.

Table 2 shows various proxies, tabulated and extended from the comprehensive study of Baker and Wurgler (2007), used to measure investor sentiment. We would like to call them “traditional” in comparison to ISV.

Table 2
Measures of Investor Sentiment

Investor Sentiment Proxy	Authors
Surveys and confidence indices	Solt and Statman (1988), Lee, Shleifer, and Thaler (1991), Swaminathan (1996), Neal and Wheatley (1998), Lee, Jiang and Indro (2002), Qiu and Welch (2005), Cliff and Brown (2005), Lemmon and Portniaguina (2006), Verma, Baklaci and Soydemir (2008)
Chat room recommendations and media factor	Antweiler and Frank (2004), Tetlock (2007)
Ambient noise level in a futures pit	Coval and Shumway (2001)
Retail investor transactions	Kumar and Lee (2006)
Over/Underreaction to earnings announcements	Barberis, Shleifer and Vishny (1998)
Stocks with extremely poor returns	DeBondt and Thaler (1985)
Mutual fund flows	Frazzini and Lamont (2005)
Trading volume	Baker and Stein (2004), Scheinkman and Xiong (2003), Barber and Odean (2008), Gervais, Kaniel, and Mingelgrin (2001), Hou, Peng, and Xiong (2008)
Dividend premia	Fama and French (2002)
New equity issues	Baker and Wurgler (2000)
Option implied volatility	Whaley (2000)
Insider trading	Seyhun (1998)
News headlines	Barber and Odean (2008) and Yuan (2008)
Advertising expense	Chemmanur and Yan (2009), Grullon, Kanatas, and Weston (2004), Lou (2008)
Price limits	Seasholes and Wu (2007)

Note: This table is tabulated and extended from the study of Baker and Wurgler (2007)

Researchers, for a long time, have attempted to explain the commonly agreed upon characteristic of financial time series data, known, as volatility clustering, through the hypothothesis that as information arrives to the market stock prices evolve.

One major traditional proxy to measure the pace of information flow in connection with the mixture of distributions hypothesis (MDH), is trading volume.

The MDH was developed by Clark (1973), Epps and Epps (1976) and Tauchen and Pitts (1983) to explain time varying volatility and volatility persistence in equity markets. According to this hypothesis, conditional time varying stock return volatility is due to a mixture of distributions, in which the stochastic mixing variable is considered to be the rate of arrival of information flow into the market. Furthermore, according to Fleming, Kirby and Ostdiek (2004), the MDH posits that return volatility is proportional to the rate of information arrival, and hence offers an explanation for the observed heteroskedasticity in returns. Thus, GARCH tests of the MDH imply that if the latent information flow variable is serially correlated, the trading volumes and return volatilities should also be serially correlated, and there should be a positive relation between them. For testing purposes, trading volume is purported as an exogenous variable in the volatility equation. What follows is that if trading volume can indeed explain volatility persistence than the GARCH parameters should be rendered insignificant and the trading volume parameter magnitudes should be significant and positively related to conditional variance. Furthermore remaining residuals should not display any serial correlation nor (G)ARCH effects.

Lamoureux and Lastrapes (1990), who use trading volume as exogenous variable to explain this phenomenon, examing the daily returns of 20 stocks from the Chicago Board of Exchange (CBOE), put forth that previous GARCH effects tend to mainly disappear upon the inclusion of trading volume. Thus, they argue that GARCH is a mere manifestation of daily time dependence in the rate of information flow to the market.

Omran and McKenzie (2000), agree with Lamoureux and Lastrapes (1990) that there is a decrease in volatility persistence with the introduction of trading volume into the variance equation. However, the authors posit that highly significant GARCH patterns remain in the squared standardized residuals for all but 4 out of 50 U.K. companies they analyze. Thus, they conclude that GARCH effects cannot be explained solely based on the serial dependence in trading volume.

Overall, there is a series of papers on GARCH models resting on the information flow-volume and stock return volatility relationship including Bollerslev and Jubinski (1999), Liesenfeld (2001), Lobato and Velasco (2000), and Girard and Biswas (2007).

These subsequent papers to Lamoureux and Lastrapes (1990), while confirming positive correlations between trading volume and volatility, present less drastic evidence of disappearance, or a dramatic reduction of GARCH effects consistent with MDH, through inclusion of trading volume in stock return volatility equations. There are also studies like Girma and Mougue (2002) that seem to be in agreement that inclusion of trading volume leads to a huge drop in volatility persistence.

While much of the literature on the U.S. (Lamoureux and Lastrapes, 1990; Kim and Kon, 1994; Gallo and Pacini, 2000), the U.K. (Omran and McKenzie, 2000) and the Australian stock markets (Brailsford, 1996) supports the linkage with the mixing variable eliminating GARCH effects, there's literature on emerging markets and markets in transition (Canarella and Pollard, 2011) that does not provide affirmative support. As suggested by Rao (2008), volatility characteristics may be attributable to the markets' own characteristics.

One such case is that of the Turkish market, where Baklaci et. al. (2011), conduct an original GARCH and probit analysis study to determine the response of firm-level stock price variation to public news and trading volume using intraday data. Among their various results, one is particularly noteworthy in the realm of the present study. The authors, determine that volatility persistence estimates do not entirely disappear through the inclusion of a news dummy and test whether trading volume as exogenous variable in the variance equation provides additional information. Their findings show that trading volume for almost all stocks, regardless of bullish or bearish markets, is positive and statistically significant, and thus, contributes significantly to explaining GARCH effects. However, they argue that while inclusion of trading volume has no dramatic effect in the reduction of volatility persistence for most stocks, it on the contrary, leads to an increase of such. In their study on the Turkish market, the news dummy appears to be a better proxy for information arrival than trading volume.

Literature on the volume volatility relationship encompasses both, individual stock-level analyses and market-level analyses, the latter reporting much weaker results in the realm of the ideas of Lamoureux and Lastrapes. Thus, it is argued that trading volume presents itself as a relatively better proxy for stock-level analysis (Gursoy, Yuksel and Yuksel, 2008: 200).

We concur with Baklacı et. al. (2011) that investors base their trading decisions on both, information arrivals in the market and beliefs and sentiments about news announcements, and that trading volume covers private information and possible noise not fully justified by public news. Thus, we include trading volume into our analysis.

2.6. ISV as New Proxy for Investor Sentiment

While the study of trading volume as an explanatory variable to the volatility of stock returns is common in literature, with the exact extent of its effect still open for discussion, ISV presents itself as a novel proxy of investor sentiment.

It not only provides insight into one of the long-studied issues in finance, but also, gives information on which region of the world, how many times, using what keywords the search is initiated. As such, it differentiates itself from the other proxies which are, more or less, ex post measures of investor actions or suffer from self-reporting biases. ISV, on the other hand, provides ex ante information on the investors thoughts but also on their inclinations, which most possibly result in actions. Furthermore, there is a cyclical effect, the more a keyword is punched in, the more likely it is to rank higher and appear more often.

ISV data was first made public in 2006 by America Online (AOL) with the intention of serving the academia. However, increased privacy concerns raised by the users led to AOL's removal of the database.

In 2008, Google began offering a free service called Google Insights, through which the tracking of keywords was made publicly available. In 2012, Google Insights was converted into Google Trends with a new interface. A seminal study by Ginsberg et al. (2008) used Google Trends data as proxy for health information seeking behaviour to track influenza illness in different regions in the U.S.

The first article with economic content using ISV data is that by Askitas and Zimmermann (2009), on unemployment figures in Germany. In this article the researchers relate search firms to job search activity to explain unemployment. Subsequent studies using unemployment figures focus on the U.S. (Choi and Varian, 2009 and, D'Amuri and Marcucci, 2009), Israel (Suhoy, 2009) and Italy (D'Amuri, 2010). All studies are in favour of using ISV and find significant evidence that this new measure can be very useful in predicting unemployment figures in a timely manner. Other studies focus on consumption related search terms to predict consumption indicators such as the study on Germany (Vosen and Schmidt, 2011) documenting that ISV data is a better proxy than traditional measures.

Da, Engelberg and Gao (2010), arguing that investor sentiment can be directly measured through the internet search behavior of households, construct a Financial and Economic Attitudes Revealed by Search (FEARS) index compiling search queries based on keywords like like recession, bankruptcy and unemployment from millions of the U.S. households. They show that the FEARS index is able to predict daily realized volatilities of ETF's even when effects of variables like volume, turnover, the VIX index and alternative sentiment measures are accounted for. In that respect their findings support the Delong et. al. (1990) noise trader model.

Preis, Moat and Stanley (2012) analyze the performance of a set of 98 search terms. Their findings show consistency that notable decreases in financial markets are preceded by periods of investor concern. They suggest that ISV information could be used to construct profitable trading strategies since investors may search for more information about the market, before deciding to buy or sell.

In a similar fashion, the rise of social and collaborative platforms such as Twitter can contribute to give insight on search behavior. For instance, Bollen, Mao and Zeng (2011) analyze the daily text content of Twitter feeds, and group them according to Google-Profile of Mood States (GPOMS) and their relation to the Dow Jones Industrial Average (DJIA). Using Granger Causality tests and Self-Organizing Fuzzy Neural network they find that the accuracy of DJIA predictions can be significantly improved through inclusion of specific public mood dimensions.

Da, Engelberg and Gao (2011) compare traditional measures and ISV and emphasize that although market-based measures have the advantage of being readily available at a relatively high frequency, they are not the equilibrium outcome of many economic forces other than investor sentiment. Also, compared to survey-based measures, search based measures have several advantages: First, search-based sentiment measures are available at a high frequency. Survey measures are often available monthly or quarterly. Furthermore, search-based measures reveal attitudes rather than inquire about them and contain inherent biases

Schmidt and Vosen (2011), comparatively study leading traditional survey-based indicators on private consumption and internet search data provided by Google Trends. They draw an interesting line between macroeconomic variables and survey-based indicators. They argue that the former indicate consumers' ability to spend, while the latter try to capture consumers' willingness to spend. They go on to say that ISV intends to provide a measure for consumers' preparatory steps to spend.

All these platforms serve as media that might demonstrate "what we collectively think" and "what might happen in the future" (Rangaswamy, Giles, and Seres, 2009: 58).

Apart from economic studies, researchers have begun to use ISV data based on names or tickers of stock market indices and individual stocks.

Bank, Larch and Peter (2011) use a multivariate panel regression model to investigate the influence of search volume on stocks listed on the German stock market index, Xetra, between the period January 2004-June 2010. They use ISV data obtained from Google Insights, a previous version of Google Trends. ISV for firm names is taken as a proxy of investor attention and investigated in terms of the impact on trading activity, liquidity and returns. The authors attribute their findings to uninformed investors and show that an increase in ISV is associated with a rise in trading activity, stock liquidity and temporarily higher future returns.

Da, Engelberg and Gao (2011) perform a similar research based on Google Trends ISV using company stock tickers for all Russell 3000 stocks between 2004-2008. The authors use the VAR model and panel regression to show that ISV is correlated with, but different from, existing proxies of investor attention such as turnover, extreme returns, news and advertising expense. In as such they determine that Internet search volume measures attention more timely than do other well-established attention variables. The authors conclude that an increase in ISV predicts higher stocks prices in the upcoming two weeks and an eventual price reversal within the year. Furthermore, ISV is found to contribute to the large first-day return and long-run underperformance for a sample of IPO stocks. Da, Engelberg and Gao (2011) reinforce the explanation that ISV measures public interest, which according to Barber and Odean (2008), implies buying pressure by uninformed retail investors in the short run.

As opposed to the previous two studies who use stock-level data, Dimpfl and Jank (2011) investigate the performance of the DJIA, FTSE100, CAC40, and DAX market indices encompossing the period of July 2006-June 2011 using VAR models and Granger causality analysis. In line with arguments of Foucault, Sraer and Thesmar (2011), the authors demonstrate that investors' attention to the stock market as measured by Google Trends ISV, rises during periods of high market movements. Furthermore, they argue that a rise in investors' attention as proxied by name-based keywords is followed by higher volatility.

Another study confirming findings of Da, Engelberg and Gao (2011), using a sample of S&P 500 stocks and their respective ticker-based keywords from Google Insights, is that by Joseph, Wintoki, and Zhang (2011). The authors, through a regression methodology, argue that in the three year period between 2005-2008 ISV, over a weekly horizon, predicts abnormal stock returns and abnormal trading volumes. Furthermore they conclude that ISV is positively linked to the difficulty of a stock being arbitraged.

Vlastakis and Markellos (2012) investigate 30 of the largest stocks traded on the NASDAQ and NYSE and their name-based queries obtained from Google Trends between January 2007-October 2009. The authors analyze the relationship between information supply, as proxied by the Reuters News Scope Archive, and, ISV data from Google Trends, which they consider as proxy for information demand. Employing correlation and causality analyses, they determine that both variables are linked contemporaneously and dynamically. Among other findings, they show that inclusion of both variables results in a significant reduction of volatility persistence by roughly 58% using a simple market model mean specification with a GARCH(1,1) model. ISV based on company names is found to be a significant regressor for 13 out 30 stocks with its sign being either a positive or negative.

Using daily trading volumes of stocks listed on the NASDAQ-100, Bordino et. al. (2012) correlate daily ticker-based ISV obtained from the Yahoo search engine arguing that such represents the attractiveness of trading of a stock. The authors apply time-lagged cross correlation and Granger causality analyses. Results of the correlation analyses indicates that ISV tends to anticipate trading volumes up to a maximum of three days and, establish that, beyond this time frame the correlation between the two variables disappears. Furthermore, their results show that this correlation only emerges at a daily scale and seems to disappear at weekly resolution. Secondly, Bordino et. al (2012) find a significant lagged cross-correlation between a volatility proxy (the absolute value of price returns) and ISV. As for Granger causality, their findings suggest that query volumes observed today have informative content of tomorrows trading volumes. Lastly, the authors track the individual user activity of individuals who have a Yahoo Profile. Findings indicate that most users search only one ticker, within a month as well as within the whole year. Thus, the authors conclude that most users of the Yahoo search engine are not financial experts.

Latoeiro, Ramos and Veiga (2013), analyze a sample of 36 companies listed on the EURO STOXX 50 Index comprising the largest companies in the Euro area. Their time frame encompasses the period of January 2004-June 2010. Weekly ISV data is collected from Google Insights. To capture abnormal variations of investor attention, the authors construct an abnormal ISV measure from name-based queries comparing current web searches to the average of the previous four weeks. They also construct an abnormal trading volume variable as in Barber and Odean (2008), and, an abnormal returns variable. Conditional volatility measures are obtained using GARCH(1,1) along with a simple market model mean specification. These variables, in addition to a realized volatility variable, form the dependent variables of their study and are sought to be determined through the abnormal ISV variable along with several control and dummy variables through regression analysis. Their results show that an increase in search queries leads to a short-lived increase in volume and volatility, which is rapidly reversed in the following week. The authors attribute the fact that the impact is higher in the following week to the presence of less sophisticated investors. Futher results indicate that web search for the market index precedes a decrease in the returns of the index and of the stock index futures, and an increase in implied volatility and, thus, is argued to show that information is not impounded by market professionals.

Our study distinctly differentiates itself from the above studies in terms of time frame, scope and data, sample construction, methodology, and the empirical justification of using name-based search queries.

The time period of this study is the broadest used so far in ISV studies encompassing the period from January 2004-September 2013. Not only do we analyze companies of various sizes belonging to the NYSE and NASDAQ composite indices, but also add another sample of Turkish firms belonging to the BIST-100 Index. Through the use of both US composite indices we guarantee the inclusion of not only large-cap companies, which is different from the other ISV studies.

Furthermore, up to the present time, all relevant studies use USD and EURO denominated stocks. The usage of Turkish stocks representing an emerging market is one of the original aspects of this study. Moreover, conducting an comparative analyes of emerging markets with developed markets, is also unprecendented.

While studies such as Da, Engelberg and Gao (2011), Joseph, Wintoki, and Zhang (2011) and, Bordino et. al. (2012) use ticker-based search queries, Vlastakis and Markellos (2012) and Latoeiro, Ramos and Veiga (2013) use name-based search queries. We concur with the latter two groups of authors and use name-based ISV data. However, different from them, we construct a control group of e-businesses to justify that ISV data, that has a significant volatility effect on non-e-businesses, actually pertains to individual investors rather than consumers.

Similarly, our meticulous stepwise procedure to arrive at our final sample of analysis, using a combination of eye-ball tests and objective criteria, is unique.

Different from previous studies, this study is purely concentrated on the phenomenon of volatility. Only the study of Vlastakis and Markellos (2012) and Bordino et. al. (2012) touch upon stock return volatility slightly in the course of their analysis, and Latoeiro, Ramos and Veiga (2013) use it in conjunction with various other variables. However, we, not only differentiate ourselves in terms of our sole focus on stock return volatility, but also, in our usage of alternative mean specifications and increasingly nested conditional volatility equations. In that regard, to the best of our knowlege, there is also no study that compares and contrasts model results for alternative mean specifications such as simple market and extended autoregressive models.

While, the inclusion of an information proxy is conceptually similar to that of Vlastakis and Markellos (2012), we, in contrast, use trading volume as opposed to a particular news archive.

While Granger causality is a common test used in ISV studies, there is no explicit mention as to the lag specifications and determination criteria used. As an added contribution, we perform VAR analyses on a company basis and determine each lag separately through the Akaike Information Criterion to be used for Granger causality analysis purposes.

In sum, this study fills the gaps of and contributes in various aspects to the few seminal studies using ISV data with respect to financial markets.

2.7. Research Questions

This chapter has dealt with what we would like to call “The Volatility Enigma” which consists of various themes adressed in literature and surrounded by alternative behavioral explanations to volatility.

Leaving aside systematic risk, it is not clear what the residual risk is truly composed of. However, there’s one factor affecting the movements of stock prices that almost every study agrees on: Information and investor sentiment.

Figure 1 illustrates the various themes surrounding volatility literature, the most relevant of which we have discussed in this chapter. It shows that information, be it public, private or historical may be picked up by a player in the stock market. The emphasis in literature is on the investor who may mistake noise for information and act upon it causing noise as explained in the Noise Trader Model. Information structure refers to an important feature that too, may influence investor behavior: the autocomplete function as demonstrated in the Appendix, which can be resembled to the framing heuristic. Before acting upon the obtained information, several heuristics may be at play leading the individual investor to cause noise. A novel proxy of investor sentiment is used in combination with trading volume to explain their mutual and ISV's isolated contribution to conditional volatility.

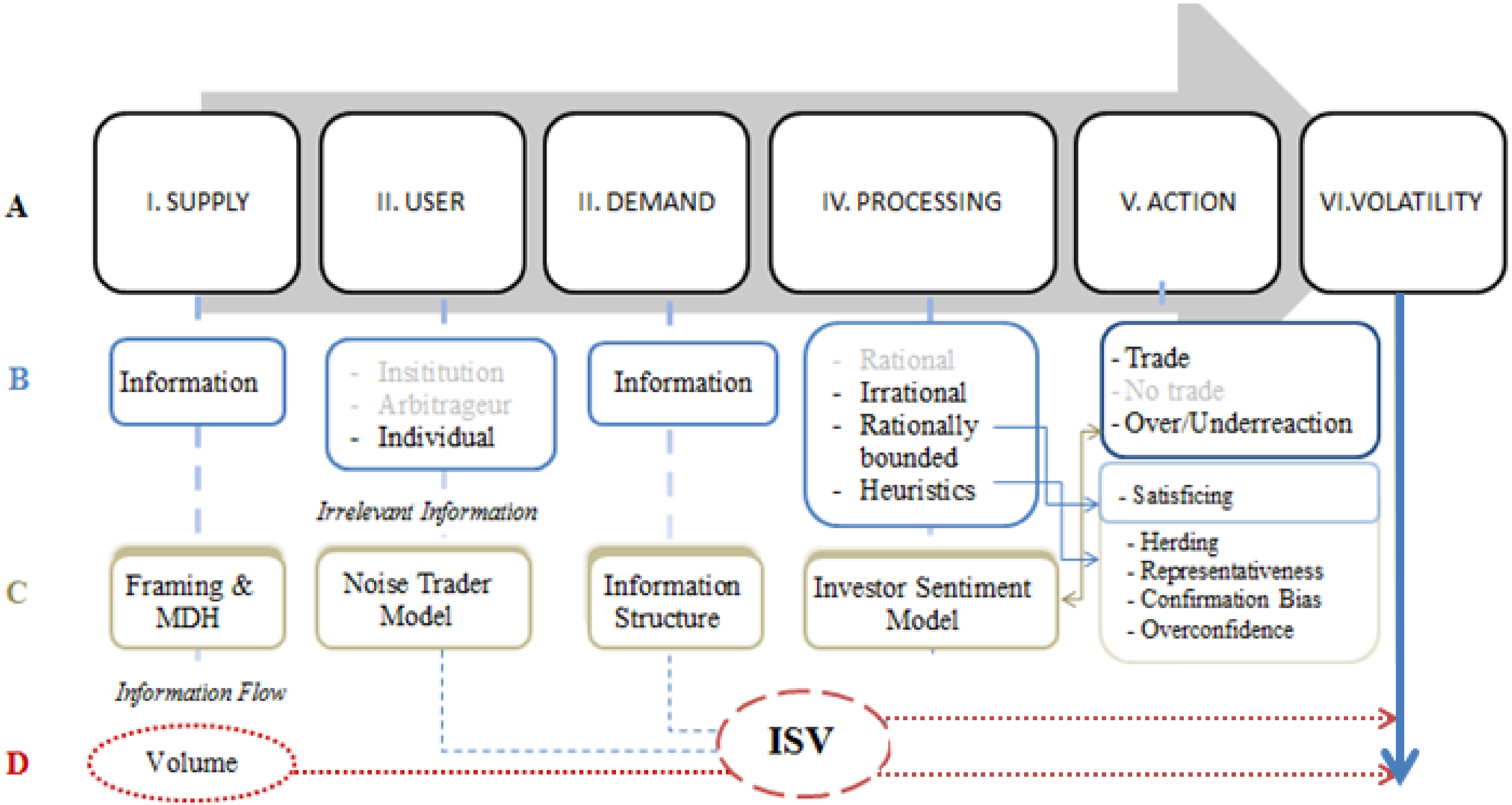

Figure 1 The Volatility Enigma

Richard Thaler wrote an article in 1999, discussing whether the end of behavioral finance was near. He made a good case for why behavioral finance cannot be dismissed since there is, both, empirical evidence and theoretical support that classical finance theories cannot entirely address. The final point of his wishlist is related to more data on individual investors becoming available. It is this point that can be adressed through the availability of ISV data.

This established theoretical framework and previous empirical findings lead us to the formulation of our research questions.

The contraversies surrounding noise trading, trading volume and the impact of investor sentiment on idiosyncratic risk present itself as our first research question:

Research Question 1 (RQ1): Does investor sentiment, as proxied by ISV, affect stock return volatility?

Grounded in the previous information flow-trading volume literature, which is more lenient towards the idea that trading volume seems to be a relatively better exogenous variable in explaining GARCH effects in developed markets relative to emerging markets, we are interested in how these two variables together, exert influence on stock return volatility. Thereby we attempt to concentrate on whether trading volume and /or internet search volume are accountable for the volatility clustering effects or GARCH effects. Thus, the second research question becomes:

Research Question 2 (RQ2): Do ISV and trading volume have any significant effect on stock return volatility?

Apart from an effect on conditional variance, we are also interested in whether there is a temporal causal linkage among stock returns, trading volume and ISV through our third research question:

Research Question 3 (RQ3): Is there a causal relationship of stock returns with ISV, and, trading volume and ISV?

We have discussed that a bulk of literature starting with Lamoureux and Lastrapes (1990) posits that inclusion of trading volume in the variance equation leads to decreases in volatility persistence, especially supported for developed markets. To this end, we formulate our fourth research question:

Research Question 4 (RQ4): Does inclusion of ISV and trading volume impact volatility persistence?

The above research questions may, as put forth by literature in the case of trading volume, yield different results for various markets. Hence, Turkish and the U.S. markets are analyzed separately to be able to pin-point potential differences in stock price variation behavior.

From a very different angle, there's still the issue of how search queries can be attributed to the individual investor versus the consumer. Thus, if a company is an e-business, in the sense that it does cater to the individual consumer through its corporate website, and if there's no significant relation of its respective search queries to its stock return volatility, we can potentially infer that search queries exhibiting significant effects for non-e-businesses belong to investors. Along these lines we form two different groups of the U.S. companies; e-businesses and non-e-businesses, and set out to infer the correctness of this assumption if e-businesses' stock price variation is not affected by the presence of the ISV variable in the variance equation. Hence, our next research question becomes:

Research Question 5 (RQ5): Can e-businesses be isolated and studied separately to establish that name-based ISV data represents individual investors rather than consumers in the case of non-e-businesses?

These five research questions form the backbone of this study and will subsequently be formulated as research hypotheses. To contribute to the global literature and refrain from taking a solely U.S. markets-based perspective we analyze stocks belonging to both, the U.S. and Turkish markets in a comparative way.

CHAPTER 3
METHODOLOGY

In this section we briefly discuss characteristics of financial time series data and relevant testing along with econometric modelling. EViews 7 is used for all modelling and analysis purposes.

The ordinary least squares (OLS) regression model that has been for long considered crucial in applied econometrics is used to determine the variation of a dependent variable in response to a change in another variable(s) called independent or explanatory. When we fit a model, the difference between the predicted value and the actual value is called the error term, or residual and denoted by the Greek letter epsilon (ε). It is this error term and its magnitude that we are interested in analyzing. The OLS model assumes that the random error term is not dependent upon the value of Xs and has a constant variance, and standart deviations over time (referred to as *homoskedasticity*). Homoskedasticity graphically indicates a uniform dispersion of data points around the regression line. However, stock prices are empirically demonstrated to be lacking homoskedasticity, they display heteroskedasticity. Heteroskedasticity is marked by non-constant measures of dispersion of the error term. Meaning that some periods are more volatile than others resulting in increased magnitude of the error term. Analyzing stock prices, we can also see that the volatility spikes are not randomly distributed over time but exhibit auto-correlation. This phenomenon is called “volatility clustering” and is explained in Section 4.1.

Robert Engle was the first researcher to study regression analysis under heteroskedasticity and developed the ARCH model in 1982.

3.1. Financial Time Series Data

A time series is a collection of data obtained by observing a response variable at periodic points in time. Alternately, if repeated observations on a variable produce a time series, the variable is called a time series variable since the data is dependent on some time increment.

Several potential difficulties may arise when working with time series data and models that use such. Limited or low frequency data may pose a problem. Another econometric problem, that limits the use of financial models is the *nonstationarity* of time series data. Nonstationary data does not have the same statistical properties (such as mean and variance) over time. A Fibonocci sequence for instance is an example of such, where at every step the sequence takes on a higher mean value. Since most econometric financial models require *stationary data*, that is data reverting back to its mean, the non-stationary data needs to be made stationry. One way to make data stationary is to take its first-differences. Thus, the data series $x_t, x_{t-1}, x_{t-2}, x_{t-3}$.... becomes $(x_t - x_{t-1}),(x_{t-1} - x_{t-2}), (x_{t-2} - x_{t-3})$. If the data is still not stationary, the second-difference can be taken. Each procedure of taking differences, however, reduces the data points in the series by one. Another method to transform nonstationary data to stationary is by means of taking logarithms.

Financial time series data commonly found to exhibit certain characteristics. One of such is the concept of "volatility clustering" going back to Mandelbrot (1963) and Fama (1965) and is marked by large (small) stock price changes are followed by large (small) stock price changes over an extended period of time. Brooks (2008) explains this almost-universal phenomenon with regard to stock prices as potentially being due to information arrivals that occur in clusters themselves than being evenly spaced through time. An outcome of volatility clustering is the observation that volatility shocks happening today affect many expected volatility periods in the future. If this is the case volatility is considered to be persistent if today's return has a large impact on the predicted variance many periods into the future (Engle & Patton, 2001: 239). Another concept traced to financial time series data, in particular, is "mean reversion", which suggests that stock prices or returns will eventually revert back to their mean, whichever way this mean may be defined.

There are several pre-tests to ensure the financial time series of analysis is stationary[3], with a constant mean, variance and autocorrelation structure that is not dependent upon time.

[3] Stationarity, in this context refers to weak stationarity

3.2. Pre-Testing

3.2.1. Unit Root Testing

To test whether a time series variable is non-stationary it needs to undergo unit root testing which is dependent upon an autoregressive model. An autoregressive process is based on the premise that a time series variable depends linearly on its past values. If a unit root is present in time series data statistical inferencing using econometric financial modelling becomes difficult. If a linear stochastic process has a unit root of 1 this process is said to be non-stationary. Unit root tests such as the Augmented Dickey Fuller (ADF) Test as developed by Dickey and Fuller (1981) are used to test for the presence of unit roots as is expressed as follows:

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta Y_{t-j} + \beta t + e_t \tag{6}$$

Where, α is a constant (the drift term), t denotes the time trend, β is the coefficient on a time trend and p is the lag order of the autoregressive process. When constraints $\alpha = 0$ and $\beta = 0$ this corresponds to modelling a random walk and if only $\beta = 0$ restriction is imposed it refers to modelling a random walk with a drift. Thus, the ADF test has three versions, all of which are available in EViews.

Before applying the test, the order of *p* needs to be determined. If too small a value is used then remaining serial correlations in the errors will bias the test. In contrast if too large a value for p is assigned then the power of the test will suffer.

The two ways by which to determine p are (1) using higher order of p's and examine the t-coefficients (2) using Akaike (AIC) or Schwartz Information Criteria (SIC) developed by Akaike (1974) and Schwartz (1978) respectively.

The decision criterion for the ADF test is the t-statistic associated with the ordinary least squares estimate of γ . This is called the Dickey-Fuller t-statistic. The null and alternative hypotheses of the Augmented Dickey-Fuller t-test respectively are:

$H_0 : \gamma = 0$ (has a unit root; thus variable is non-stationary)
$H_1 : \gamma < 0$ (has no unit root; thus variable is stationary)

The decision rule is:

If *critical t value > ADF test stat* → do not reject H_0; there is a unit root
If *critical t value < ADF test stat* → reject H_0; unit root does not exist

3.2.2. Testing for ARCH Effects in the Error Term

We mentioned in the beginning of this chapter, that standart OLS models assume the residual error term of the regression equation to be with a zero mean and constant variance $\{\varepsilon_t \sim N(0,\sigma^2)\}$ and thus base their assumptions on homoskedascity.

However, stock prices are observed to display features such as volatility clustering (the data is autocorrelated), leverage effects (the inclination for volatility to rise more subsequent to a price fall than subsequent to a prise rise of the same magnitude) and leptokurtosis (fat tails and excess peakedness at the mean). Thus, ARCH models accounting for the heteroskedastic nature of errors became widespread. The autocorrelation in volatility is modelled by ARCH models in by allowing the conditional variance (σ_t^2 , or often times denoted as h_t) of the residual error term from the mean equation (which can be defined linearly or non-linearly) to depend on the immediate previous value of the squared error.

This conditional mean, from which the conditional error variance equation is drived, depends upon the specification of the researcher. Together the conditional mean and conditional variance equations form a system that is estimated through an iteration process using maximum likelihood. The selection of an appropriate mean specification is crucial since the error term derived from that equation is what is being modelled in the variance equation. For the mean specification, we rely on previous literature and use a set of increasingly nested models. These range from a market model following Lotaeiro, Ramos and Vega (2013) to an AR(1) model along the lines of Baklacı et. al. (2011), while the latter is more commonly used and theoretically presents itself as a better choice for econometric analysis in behavioral finance. Consistent with Vlastakis and Markellos (2012), we include the market return in both mean specifications. For conditional variance modelling we adopt GARCH(1,1) based on both, popular usage in financial econometrics analysis and the lag selection procedure based on the minimum AIC.

An autoregressive model is a very common time series model. For a time series y_1, $y_{,2}$, y_3, …, y_n, an autoregressive model of the order of p, which is denoted as AR(p) states that y_t, is a linear function of previous p values of the series plus an error term:

$$y_t = \alpha + \delta_1 y_{t-1} + \delta_2 y_{t-2} + ... + \delta_p y_{t-p} + \varepsilon_t \text{, where } \varepsilon_t \sim \mathrm{N}(0,\ \sigma_t^2) \tag{7}$$

The two mean specifications applied in this study are depicted in equation (8a,b), where the former is the market model and the latter is an AR(1) model with the market (index) return as exogenous variable in the mean:

$$y_t = \alpha + \beta x_1 + \varepsilon_t, \qquad \text{where } \varepsilon_t \sim \mathrm{N}(0,\ \sigma_t^2) \tag{8a}$$

$$y_t = \alpha + \beta x_1 + \delta y_{t-1} + \varepsilon_t, \qquad \text{where } \varepsilon_t \sim \mathrm{N}(0,\ \sigma_t^2) \tag{8b}$$

Here, α is the constant, β is the parameter of the market return (market index return) x_1, δ is the parameter for y_{t-1} the previous stock return also called AR(1), and ε_t is a random error with a conditional variance. The difference between the above two mean specifications is that the second model is an AR(1) model specifying that y_t, depends linearly on its own previous value. Necessary conditions for the AR(p) model to remain wide-sense stationary are absolute value of δ must be lower than 1, the roots of the polynomial must lie within the unit circle. For AR(1) processes with positive parameters, solely the previous term and the noise contribute to the change in stock price. If the parameter is close to 0, the process looks like white noise and the more the parameter approaches unit the more is contributed from the previous term rather than the error term.

We have also tested various other mean specifications where we include trading volume and ISV variables in the mean equations interchangibly while leaving them out of the conditional variance equations one by one. Since neither ISV nor trading volume have a noteworthy statistically significant impact on the mean, they have been exluded from the mean equations. The same procedure has been applied with lags of trading volume and ISV variables, again no significant linear contribution to the mean was obtained.

Equation (9) are illustrations of ARCH(1) with the conditional variance depending on only one lagged squared error and ARCH (q) where error variance depends on q lags of squared errors:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2$$
$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots\dots\dots + \alpha_q \varepsilon_{t-q}^2 \qquad (9)$$

Where, all coefficients are required to be non-negative ($\alpha_i > 0$). In order to apply ARCH-type models, detection whether ARCH effects are present in the residual of an estimated model, is needed. This testing procedure was originally devised by Engle (1982) and is similar to the Lagrange Multiplier (LM) test for autocorrelation. Brooks (2008) explains the procedure for testing for ARCH effects for pre-testing as follows:

(i) Run any postulated linear regression
(ii) Square the residuals, regress them on q own lags (run regression):

$$\hat{\varepsilon}_t^2 = \gamma_0 + \gamma_1 \hat{\varepsilon}_{t-1}^2 + \dots\dots\dots + \gamma_q \hat{\varepsilon}_{t-q}^2 + \vartheta_t \text{, where } \vartheta_t \text{ is the error term} \qquad (10)$$

and obtain R^2.

(iii) The test statistic is defined as TR^2 (the number of observations multiplied by the coefficient of a multiple correlation) from the last regression and is distributed as χ^2 with q degrees of freedom.

(iv) The null and alternative hypotheses are:

$H_0 : \gamma_1 = 0$ and $\gamma_2 = 0$ and… $\gamma_q = 0$

H_1 : At least one of $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$ and… $\gamma_q \neq 0$

The software provides the test for detecting the presence of ARCH in residuals by regressing the squared residuals on a constant and *p* lags, where *p* is user-defined.

Another indicator of conditional heteroskedasticity is Kurtosis, which is a measure of whether the data is peaked or flat relative to the benchmark normal distribution. The kurtosis of normal distribution is 3, values above 3 (defined as *leptokurtic*, a feature generally associated with financial time series), indicate that a distribution has fatter tails and the chance of extreme outcomes is more compared to a normal distribution.

3.3. Modelling

ARCH(q) models depicted in equation (9) pose a difficulty in empirical application: they need for a large number of parameters to explain serial dependence of the variance. Instead, Bollerslev (1986) introduced the more parsimonious GARCH model as an extension to the ARCH term (lagged squared errors) by adding the GARCH term (lagged conditional variances) as explanatory variable. A GARCH(p,q) has two characteristic parameters: the number of GARCH terms defined by p referring to the number of autoregressive lags and the number of ARCH terms defined by q referring to the number of moving average lags.

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{11}$$

Equation (11) depicts a GARCH(1,1) model where there is a constant (ω), an ARCH term ($\alpha_1 \varepsilon_{t-1}^2$) at first lag and a GARCH (σ_{t-1}^2) term at first lag, with positivity constraints for $\omega > 0$, *and parameters* $\alpha_1 \geq 0$, $\beta_1 > 0$, and $\alpha_1 + \beta_1 < 1$.

The GARCH (1,1) model solves for the conditional variance as a function of its previous variance, its previous squared return and the long-run variance.

The sum of the ARCH and GARCH term parameters is called *volatility persistence* and refers to how quickly the variance reverts or “decays” toward its long-run average. If persistence is high (low), this means that the decay and the reversion to the mean is slow (quick). If the sum of ARCH (α) and GARCH (β) parameters is 1, this implies there is no mean reversion. If persistence is less than 1, this means there is a reversion to the mean. If persistence is low, this implies a greater reversion to the mean. If the sum is above 1, this implies non-stationarity and the GARCH model cannot be used and replaced by for instance an Integrated GARCH (IGARCH) model (Engle and Bollerslev, 1986).

3.4. Post-Modelling (Diagnostic Tests)

To determine whether the proposed ARCH-type model is a good fit and properly, one needs to check the residuals of the conditional heteroskedasticity equation. Commonly applied tests to residuals of conditional volatility models are as follows:

3.4.1. Testing for Serial Correlation (Ljung –Box Q Statistic)

A practical statistic to detect whether there is serial correlation is the Durbin Watson (DW) statistic that is part of the standart regression output of any statistical software package. The DW, however has its shortcomings in that it only measures first-order serial correlation (the linear relation between adjacent residuals from the regression equation). A rule-of-thumb is that if it is very close to 2, then there is no serial correlation and if it is below (above) 2, then there’s a potential positive (negative) correlation.

Another test, that overcomes shortcomings of the DW statistic is the Ljung - Box *Q* statistic for higher order serial correlation detection. Defined as:

$$Q_{LB} = T(T+2)\sum_{k=1}^{s} r_k^2 / (T-k) \tag{12}$$

Where, T is the sample size, r_k^2 is the sample autocorrelation at lag k and s is the number of lags being tested. If the test statistic is bigger than the chi-squared distribution with s degrees of freedom set at a certain significance level then we reject the null hypothesis.

The null hypothesis of this test is that there is no serial correlation in the residuals up to a specified order (s) and specified as:

$H_0 : r_k = 0$ and k=1,..s (the data is independently distributed)

$H_1 : r_k \neq 0$ for at least one k=1,..s (the data displays serial correlation), where r_k is the k-th autocorrelation.

Similarly, p-values that are significantly larger than 0, are considered as an indicator that the null hypothesis; that there is no serial correlation left in the residuals, can not be rejected.

3.4.2. Testing for ARCH Effects

Heteroskedasticity tests (Engle's LM) as outlined above for the residuals of the GARCH equations are also administered post-modelling. If the Chi-Square p value is significantly large this means that previous ARCH effects are captured by our model and none is left.

3.5. Granger Causality Testing for Causal Ordering

Granger causality (Granger, 1969) is a statistical hypothesis tests that aims to identify causality through lead-lag relationships based on the premise that correlation does not imply causation. By examining variables in time based on which one is occuring prior to the other, Granger causality may be established. One variable does not Granger-cause the other, if adding past observations of the former to the information set with which we forecast the latter does not improve this forecast. In other words, a time series X is said to Granger-cause Y if it can be shown, usually through a series of t-tests and F- tests on lagged values of X (and with lagged values of Y also included), that those X values provide statistically significant information about future values of Y.

The methodology of establishing Granger causality is fairly simple and is available in EViews. If a time series is stationary, the test is performed using the level values of two (or more) variables. The number of lags is chosen based on minimum AIC or SIC. Any particular lagged value of one of the variables is kept in the regression if (1) it is significant according to a t-test, and (2) it and the other lagged values of the variable jointly add explanatory power to the model according to an F-test. Then the null hypothesis of no Granger causality is not rejected if and only if *no lagged values* of an explanatory variable have been retained in the regression.

CHAPTER 4
THE MODEL, DATA AND EMPRICAL FINDINGS

Chapter 1 and 2 have presented the relevant theoretical framework and derived the research questions of this theses, followed by an account of econometric financial models that are typically used in conditional volatility studies.

The first part of this chapter aims at formulating the hypotheses and translating them into such empirically testable models. The second part addresses the topic of data and sampling. Lastly empirical findings are presented.

4.1. Hypothesis Formulation

The derivation of the research hypotheses from the research questions, formulated at the end of the second chapter, is given in Table 3:

Table 3
Overview of Research Hypotheses

RQ	Hypothesis	
1	H_1	There is a significant ISV impact on conditional volatility
2	H_2	ISV and trading volume significantly affect conditional volatility
3	H_{3a}	There is a Granger causal relationship of stock returns with ISV
	H_{3b}	There is a Granger causal relationship of trading volume with ISV
4	H_4	ISV and trading volume contribute to a decrease in volatility persistence
5	H_5	E-businesses can be used as control groups to justify usage of name-based queries

Note: Hypotheses (H) and the relevant reference number of the empirically tested models (M) that pertain the research questions (RQ), GC denoters Granger Causality testing.

There are five research hypotheses: The first one posits that ISV has a significant effect on stock return volatility. The second one differentiates itself from the first research hypothesis in that it conjectures the effect of the two exgenous variables, ISV and trading volume, on stock return volatility. The third research hypothesis posits the existence of a Granger causal relationship. The fourth research hypothesis states ISV and trading volume have a significant effect in decreasing volatility persistence. The fifth research hypothesis is related to the e-business group and posited to justify the use of name-based search queries as being representative of individual investors in the case of non-e-businesses.

Details of the models applied are shown in Table 4, which presents these research hypotheses as equations.

Table 4
Model Specifications

Model	Specification	Conditional Mean	GARCH-Conditional Variance
1	MM-G(1,1)	$r_t = c + \lambda M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i h_{t-i}$
2	MM-G(1,1)-ISV	$r_t = c + \lambda M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i \sigma_{t-i}^2 + \psi_{ci} ISV_t$
3	MM-G(1,1)-ISV-V	$r_t = c + \lambda M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i \sigma_{t-i}^2 + \psi_{ci} ISV_t + \psi_{vi} V_t$
4	AR(1)M-G(1,1)	$r_t = c + \delta r_{t-1} + \lambda_1 M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i h_{t-i}$
5	AR(1)M-G(1,1)-ISV	$r_t = c + \delta r_{t-1} + \lambda_1 M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i \sigma_{t-i}^2 + \psi_{ci} ISV_t$
6	AR(1)M-G(1,1)-ISV-V	$r_t = c + \delta r_{t-1} + \lambda_1 M + \varepsilon_t$	$\sigma_t^2 = \omega + \alpha_i \varepsilon^2_{t-i} + \beta_i \sigma_{t-i}^2 + \psi_{ci} ISV_t + \psi_{vi} V_t$

Note: (1): r_t is expected conditional stock return (2) c is the constant (3) ε_t is residual returns (4) λ and δ are the parameters for market (index) return and previous own value of r_t in the mean equation (5) ω is the constant (or unconditional variance term) (6) α_i is the parameter for the ARCH term (7) ε^2_{t-i} is news about volatility from the previous period, measured as the lag of the squared residual from the mean equation (the ARCH term) (8) β_i is the parameter for the GARCH term (9) σ_{t-1}^2 is last period's forecast variance (the GARCH term) (10) ψ_c, ψ_v are the parameters for ISV and trading volume, respectively. (11) G stands for GARCH, AR stands for autoregressive at order of 1 and M stands for market (index) returns.

Table 4 depicts the mean and variance specifications used in this study in an increasingly nested manner. Models one through three are the basic market base model, the basic market model with ISV in its conditional variance and the basic market model with ISV and trading volume in its conditional variance, respectively.

Models four through six differentiate themselves from the previous three models in that their mean equation includes the additional AR(1) term.

4.2. Assumptions

The major assumption underlying this study is that, ISV data supplied by Google Trends can used as adequate a proxy for Turkish investor sentiment as it can be used for US investor sentiment. Furthermore we assume that Google ISV data is representative of all search queries in its database regardless from which device, be it a smart phone or laptop, the queries are initiated. Lastly, structural breaks in the economy can be disregarded since this study investigates whether there is an investor sentiment effect on volatility for stocks in the same time period and not whether the effect of investor sentiment changes between time periods. Thus, there is no necessity to account for structural breaks. Secondly, if we viewed the Internet Bubble in the end of 1999 and the mortgage crises of 2008 as two major structural breaks, we would not have data (Google ISV data is available only from 2004 onwards) to study the former or not have *enough* data points available (Google ISV data is available in weekly frequency) to perform an intuitive time series analysis,

4.3. Sampling

Two main groups of companies are used in this study: the U.S. companies listed on NYSE and NASDAQ composite indices and Turkish companies listed on BIST-100 index. The U.S. companies are analyzed in groups of e-businesses and non-e-businesses. The determination of the companies of analysis is explained step by step below:

The initial step involves the determination of the components of the three indices: For the U.S. market they are the NASDAQ Composite (ticker: IXIC) and the

NYSE Composite (ticker: NYA) with *2476* and *1867* firms listed, and cross-listed at the time of analysis, respectively. With respect to the Turkish market it is the *100* companies listed on the BIST-100 index. The reason of this choice of these three popular indices rests on the logic that the more known a company is the higher the chances of obtained complete ISV data from Google Trends.

The second step is an eyeball tests for each company from these three lists in order to eliminate companies that have more than two, preferably only one, name and/or have a generic name (ie. American Campus Communities, Inc; Qihoo 360 Technology Co. Ltd) because Google Trends results are mostly non-existing for unpopular companies, and if present, not pertaining to the companies. Some of the few exceptions to this rule are very prominent companies like Apple. Since Google Trends data starts in 2004, almost all companies who had their IPOs later than 2004 are eliminated to focus on the companies who have the maximum number of data points in their time series variable. From this elimination a total of *221* Turkish and US companies remain. We search the name for each company, to determine whether search results pertain to the respective company, by cross checking the news headlines feature displayed on their ISV index graphs as shown in the Appendix, subsequently we download ISV data from Google Trends and check whether it is in a format fit for analysis since low-volume data contains many irrelevant “0” values over long time periods. As a result, a total of *102* companies remain.

The third step involves “interim econometric analysis” through data transformation into Log Price Returns and Log ISV by taking the logarithms of the change in price “Log (P_t/P_{t-1})”, ISV data “Log (ISV_t/ISV_{t-1}), and trading volume “Log (V_t/V_{t-1}), a common procedure for stock return volatility analysis. At this step pre-testing for ARCH effects for each company is performed as well. Out of 102 US and Turkish companies *81* are found to have ARCH effects in their residuals, and hence could undergo GARCH model testing.

The fourth step involves determining the e-business firms out of the 73 remaining U.S. companies. Hence, 73 companies’ websites are analyzed to check for the presence of e-commerce (ie. the availability of a virtual shopping card option is available on corporate website) and, in case of the availability of such, the firms are cross-checked with the Internet Retailers 500 List of E-Businesses. Resultingly, *26* are determined to be e-businesses.

The fifth step involves “interim financial modelling” by applying to these 81 companies the Model M1 with subsequent diagnostic testing for remaining serial correlation or heteroskedasticity.

Out of 73 US companies, 27 are unusable for further analysis since they fail diagnostic testing. At the end of this step *46* US companies (34 are non-e- businesses and 12 are e-businesses) and 8 Turkish companies remain amounting to a total of *54* companies. With regard to e-business companies, out of the 12 remaining companies we exluded Dell and Microsoft because although they are listed in the Internet Retailers 500 List of E-Businesses, interviews with local executives and analysis of their annual reports shows that they derive most of their revenue not through online sales but rather through physical stores.

The next step involves checking whether all *52* companies can be analyzed under the GARCH modelling framework recalling the positivity constraints imposed upon the conditional volatility parameters. While all 52 companies are analyzed using all models, it is crucial to note that five U.S. non-e-business, two U.S. e-business and two Turkish companies do not fulfill this criterion in the base GARCH models. All results are reported nevertheless. To remedy such shortcomings various other ARCH or GARCH lag specifications or the EGARCH model (Nelson, 1991) could be used in further studies, are, however, beyond the scope of this thesis.

A shortcoming of Turkish firms is that most do not have meaningful Google Trends data available and, among the ones that do, the data is not available starting 2004. As a result, eight Turkish companies remained.

4.4. Data and Descriptive Statistics

Stock return and trading volume data for the U.S. and Turkish companies is obtained on a weekly basis from Yahoo Finance and Reuters, respectively. US stock returns are based on US Dollars and Turkish stock returns are based on Turkish Lira.

The ISV data collection procedure is as follows: The Google Trends website provides a search box where any choice of keywords can be typed in. For meaningful results Google Trends makes available an index representing search intensity.[4] This indexed search volume data is available on a weekly format covering the same period as the stock return data. Another important remark on Google Trends data is that due to the sampling method Google uses the results vary from observation to observation and only search queries above a certain volume are being included into the query index.

As discussed, the few studies that use ISV data, are divided between what represents investor sentiment better; the firm name or ticker symbol. We have obtained data for both, the name, the ticker symbol and in addition (which none of the cited studies has considered) we obtained data based on world-wide search queries and only US-based search queries.

Data is downloaded for free with a Google Account in ".cvs" format and converted into Excel. In most cases, world and US-based search queries show significant correlations however we chose to use world-wide search queries since we are interested in not only the sentiment of investors based in the U.S. but also the universal investor sentiment. The name-queries and ticker-queries of most companies show some significant correlation as well, however name-based search queries generated longer and more company-relevant time series. For instance as presented by Da, Engelberg and Gao (2011), if investors are searching "AAPL" (the ticker for Apple Computer Inc.) in Google, they are likely to be interested in financial information about the stock of Apple Inc.. Thus, we use name-based ISV for all stocks and indices, and "Istanbul Stock Exchange (ISE) Index" the latter being the old name for Borsa Istanbul.

In contrast to the U.S. where we use world-wide ISV data, in the Turkish case only Turkish regional search queries are considered. This is because, firstly, there is relatively less-to-none relevant global ISV data available on Turkish companies. For instance when you do a global Google Trends search for the word "DESA" you

[4] For methodology of index construction visit www.google.com/trends

obtain results mainly belonging to the region Sri Lanka that do not pertain to to the Turkish company. And secondly, using regional Turkey ISV data makes intuitively more sense since foreign-based investors looking to invest into the Turkish market can be considered “sophisticated” and use institutional managers and do not consult search engines on whether to buy particular Turkish stocks.

We transform all data into logarithmic series and implement ADF unit root test for all three options: with intercept, with intercept and time trend, with neither an intercept nor a time trend. None, as is expected after applying this transformation process, displays unit roots afterwards.

The descriptive statistics are given in Table 5. As is reported, kurtosis and skewness statistics, also jointly represented by the J-B statistic, show that the price data is not normally distributed. NYSE companies deviate more from normality than NASDAQ companies and drastically more than Turkish companies.

Table 5
Descriptive Statistics for Stock Price Returns

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AXE	0,27	0,37	29,75	-41,88	5,17	-0,72	14,74	2933,08	0,00
BHI	0,07	0,27	17,20	-40,82	5,56	-1,10	10,18	1177,65	0,00
BMI	0,18	0,48	32,39	-37,49	6,16	-0,33	9,43	620,94	0,00
BMS	0,15	0,15	11,17	-12,78	3,14	-0,09	4,58	52,84	0,00
CIR	0,45	0,61	14,23	-20,98	5,93	-0,52	3,99	20,43	0,00
CVX	0,27	0,54	15,47	-31,67	3,50	-1,56	16,98	4285,09	0,00
DUK	0,21	0,27	13,11	-18,49	2,50	-0,70	10,79	1305,81	0,00
DVN	0,15	0,33	16,56	-31,86	4,87	-0,67	7,23	412,04	0,00
EMC	0,11	0,13	17,06	-20,94	4,34	-0,26	5,05	93,58	0,00
EME	0,25	0,38	24,06	-23,97	5,30	-0,19	6,11	205,12	0,00
F	0,03	-0,08	62,96	-71,13	7,63	-0,39	27,18	2214,91	0,00
HAL	0,27	0,32	18,26	-43,14	5,87	-1,14	9,79	1073,26	0,00
HON	0,21	0,31	13,20	-21,26	3,77	-0,61	6,27	254,52	0,00
KO	0,14	0,10	7,19	-23,67	2,49	-1,82	9,83	6188,36	0,00
ODP	-0,23	0,00	74,44	-52,82	9,79	0,09	15,09	3090,68	0,00
OMX	-0,15	0,00	87,52	-60,41	9,38	0,52	23,66	9043,51	0,00
OXY	0,31	0,29	19,53	-33,77	4,83	-0,74	9,26	863,31	0,00
PFE	0,03	0,00	13,32	-22,77	3,33	-0,81	9,65	979,37	0,00
PG	0,14	0,20	9,19	-17,60	2,29	-0,88	10,46	1226,47	0,00
SO	0,15	0,16	11,93	-14,11	2,09	-0,42	9,66	940,08	0,00
WAG	0,09	0,08	12,34	-18,58	3,60	-0,56	6,92	346,55	0,00
WMT	0,10	0,18	8,24	-15,90	2,62	-0,61	6,07	230,03	0,00
NYSE	0,04	0,17	11,18	-20,61	2,48	-1,04	14,04	2641,21	0,00
AAPL	0,76	0,98	18,09	-27,84	5,22	-0,54	5,46	150,82	0,00
ABFS	-0,10	-0,07	45,42	-23,46	6,96	0,63	7,74	456,04	0,00
ACET	0,10	0,11	21,28	-28,99	5,37	-0,21	5,92	181,88	0,00
ACTG	0,25	0,33	32,72	-58,24	8,37	-0,68	8,89	762,09	0,00
ADBE	0,18	0,28	15,11	-21,69	4,54	-0,52	5,61	164,15	0,00
ADI	0,04	0,27	19,23	-19,02	4,09	-0,14	5,28	110,49	0,00
ALTR	0,09	0,16	14,34	-20,56	4,68	-0,26	4,23	37,18	0,00
ASTE	0,19	0,30	38,43	-30,53	6,59	0,04	7,14	357,63	0,00
AXAS	0,06	-0,50	52,93	-34,66	10,13	0,73	6,61	318,35	0,00
BIIB	0,31	0,45	26,73	-58,37	5,22	-3,25	39,16	28172,96	0,00
COST	0,26	0,38	11,29	-15,67	3,17	-0,64	6,19	246,41	0,00
CTXS	0,24	0,44	17,51	-28,84	5,07	-0,49	5,79	182,44	0,00
EQIX	0,38	0,61	25,64	-40,93	6,20	-0,98	9,77	843,74	0,00
EXPE	0,07	0,07	27,42	-31,43	6,54	-0,42	7,56	378,34	0,00
FISV	0,19	0,25	13,39	-20,51	3,51	-0,56	8,13	575,43	0,00
INTC	-0,03	-0,17	16,88	-16,97	4,03	-0,10	5,57	138,79	0,00
MYL	0,08	0,05	31,41	-35,24	4,94	-0,81	14,06	2605,62	0,00
NILE	0,11	0,29	28,94	-28,92	7,24	-0,11	4,81	67,18	0,00
REGN	0,73	0,73	20,60	-29,47	7,39	-0,27	3,79	13,52	0,00
SNDK	0,12	0,09	36,12	-67,07	8,36	-0,87	13,74	2473,11	0,00
VRTX	0,39	0,20	50,55	-26,30	7,08	1,23	11,21	1533,98	0,00
XLNX	0,05	0,32	14,58	-19,59	4,22	-0,18	4,25	35,20	0,00
NASDAQ	0,14	0,31	10,46	-14,70	2,85	-0,50	5,76	180,50	0,00

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AKBNK	0,29	0,00	23,15	-21,66	5,85	-0,09	4,47	46,55	0,00
ALARK	0,14	0,00	16,54	-18,30	4,46	-0,33	4,95	88,91	0,00
DESA	-0,04	0,00	27,92	-34,60	6,48	-0,68	7,87	443,57	0,00
PETKM	0,20	0,00	22,13	-14,06	4,77	0,39	4,94	76,42	0,00
SANKO	0,00	0,00	21,64	-20,07	4,63	-0,07	6,42	203,85	0,00
THYAO	0,42	0,00	21,87	-27,51	6,00	-0,25	4,64	2,41	0,00
TOASO	0,50	0,75	24,39	-44,83	6,76	-0,82	8,67	606,08	0,00
YATAS	0,01	0,00	24,92	-43,58	7,19	-0,99	10,33	1001,83	0,00
BIST100	0,28	0,59	15,76	-19,27	3,91	-0,45	5,03	103,87	0,00

Note: SD, SK, K, J-B and p stand for standart deviation, skewness, kurtosis, Jarque-Bera statistic and its corresponding p-value, respectively. U.S. prices are based on USD Dolar and Turkish prices are based on Turkish Lira.

When US e-business vs US non-e-business statistics are compared, we find no major difference in terms of the deviation statistics from normality for both groups. The approximate mean returns for NYSE, NASDAQ and BIST-100 listed companies and (and their corresponding indices) are 0,15% (0,04%), 0,20% (0,14%) and 0,19% (0,28%). Thus, we can infer that on the average, the markets have followed a bullish trend, on the average, in this period.

For e- vs. non-e-businesses the mean returns are 0,14% and 0,19%. However, it deems important to note that mostly trading costs on NASDAQ are typically greater than on the NYSE which reduces returns and needs to be factored into comparative analyses. As for standart deviations of NYSE, NASDAQ and Turkish stocks and their corresponding indices the values are: 4,73% (2,48%), 5,86% (2,85%) and 5,77% (3,91%), suggesting that all stocks, on the average, have been more volatile than their indices during the period.

Table 6 and 7 report the descriptive statistics for the logarithmically transformed trading volume and ISV series.

Table 6
Descriptive Statistics for Trading Volume

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AXE	-0,10	-2,69	184,54	-138,73	44,10	0,43	4,03	37,44	0,00
BHI	-0,06	-0,84	151,67	-87,34	30,23	0,74	5,74	202,06	0,00
BMI	-0,01	-3,50	171,23	-176,96	49,90	0,38	3,75	23,59	0,00
BMS	-0,04	-1,71	157,28	-104,09	34,59	0,34	3,90	26,51	0,00
CIR	-0,60	-5,36	113,69	-120,46	44,00	0,12	2,98	0,56	0,76
CVX	0,00	0,00	86,65	-78,16	22,55	0,15	3,73	13,07	0,12
DUK	-0,07	-1,01	96,36	-95,68	31,40	0,05	3,44	4,32	0,12
DVN	-0,08	-1,43	128,46	-92,08	26,57	0,41	4,58	66,35	0,00
EMC	-0,12	-1,93	125,01	-89,56	35,19	0,34	3,48	14,26	0,00
EME	-0,03	-0,71	111,41	-109,41	29,57	0,20	4,37	42,65	0,00
F	0,27	-2,93	168,50	-115,94	38,07	0,48	4,31	55,28	0,00
HAL	0,02	-1,88	111,65	-106,04	30,54	0,33	3,97	29,12	0,00
HON	-0,15	-1,47	106,65	-99,38	29,14	0,24	3,92	22,64	0,00
KO	0,09	-0,89	90,05	-99,95	28,58	0,10	3,46	5,28	0,07
ODP	-0,01	0,00	37,11	-21,87	6,86	1,01	7,03	425,32	0,00
OMX	0,03	-2,42	254,32	-186,87	45,16	0,56	5,73	181,69	0,00
OXY	-0,03	-0,85	91,01	-85,02	25,33	0,10	3,56	7,30	0,03
PFE	0,05	0,32	111,28	-109,79	29,06	0,09	4,13	27,13	0,00
PG	0,03	-1,61	90,70	-81,88	28,02	0,24	3,40	8,19	0,02
SO	0,04	0,43	47,02	-105,76	30,57	0,16	4,80	69,64	0,00
WAG	0,09	-2,16	167,04	-116,75	34,65	0,36	5,35	126,45	0,00
WMT	-0,11	-1,56	117,51	-121,65	30,14	0,42	4,67	72,53	0,00
AAPL	0,05	-1,88	114,69	-91,54	32,84	0,41	3,58	20,70	0,00
ABFS	0,32	-2,03	129,85	-121,62	43,78	0,31	3,27	8,78	0,01
ACET	-0,12	-1,65	337,86	-212,10	54,95	0,68	7,17	401,55	0,00
ACTG	0,11	-0,61	207,92	-179,30	56,77	0,40	3,99	33,98	0,12
ADBE	-0,22	-0,06	175,61	-127,78	38,69	0,19	4,39	43,14	0,00
ADI	-0,15	-2,33	110,22	-105,33	31,02	0,47	3,70	28,71	0,00
ALTR	-0,20	-2,78	126,82	-106,92	33,11	0,20	3,90	20,19	0,00
ASTE	0,01	-3,17	204,90	-129,63	50,25	0,35	3,58	17,29	0,00
AXAS	0,24	-1,49	204,40	-165,62	49,61	0,31	4,09	32,87	0,00
BIIB	-0,40	-2,02	289,54	-149,98	40,01	1,06	10,05	131,47	0,00
COST	-0,20	-3,72	116,97	-84,16	33,02	0,47	3,49	23,61	0,00
CTXS	-0,41	0,59	154,22	-121,76	39,74	0,22	3,87	19,79	0,00
EQIX	0,12	-1,66	282,12	-178,00	46,88	0,52	5,82	153,84	0,00
EXPE	0,84	-1,71	288,56	-111,68	43,27	1,22	8,30	597,61	0,00
FISV	-0,31	-0,76	107,58	-103,89	32,46	0,24	3,42	8,62	0,01
INTC	-0,10	-2,05	140,13	-107,44	29,80	0,63	5,26	140,01	0,00
MYL	-0,01	-3,76	212,50	-150,76	47,37	0,67	5,09	127,91	0,00
NILE	-0,63	-1,69	297,82	-170,17	57,28	0,48	5,23	118,71	0,00
REGN	0,21	-1,72	190,16	-155,15	49,04	0,40	4,19	30,98	0,12
SNDK	-0,35	-2,78	117,51	-112,46	38,95	0,19	3,19	3,79	0,15
VRTX	0,06	-2,54	260,61	-152,22	49,15	0,54	5,34	139,13	0,12
XLNX	-0,35	-0,86	143,24	-134,33	34,86	0,11	4,19	30,45	0,00

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AKBNK	0,13	-0,96	363,00	-364,36	56,49	0,09	13,88	2511,89	0,00
ALARK	-0,37	-5,52	568,58	-500,14	78,04	0,43	12,95	2116,48	0,00
DESA	-0,03	-4,91	458,16	-372,93	86,95	0,66	6,63	258,38	0,00
PETKM	-0,05	-5,79	444,29	-410,33	74,24	0,34	10,90	1091,59	0,00
SANKO	-0,24	-8,87	363,91	-412,72	95,13	0,27	4,68	53,94	0,00
THYAO	-0,03	-0,48	375,85	-317,98	68,92	0,50	7,38	427,40	0,00
TOASO	-0,44	-1,29	453,41	-439,48	66,19	0,20	13,36	1867,69	0,00
YATAS	0,43	-4,21	532,79	-341,83	92,16	0,50	6,93	285,45	0,00

Note: SD, SK, K, J-B and p stand for standart deviation, skewness, kurtosis, Jarque-Bera statistic and its corresponding p-value, respectively.

Table 6 depicts that, trading volume for 8 out 44 US companies is close to being normally distributed with J-B p-values being above 5% significance. However, trading volume data here includes all trades, executed by all sorts of investors, be it irrational, rational, individuals or institutions. For ISV data on the other hand, we would like to underline our assumption that name-based ISV data is highly likely to represent the individual investor.

Along similar lines, standart deviations for trading volume data for NYSE, NASDAQ and Turkish companies are very high compared to ISV and stock return series, namely; 32,01%, 42,86% and 77,26%. This may be an outcome of the fact that companies with different market capitalizations are represented here. For instance, although both companies belong to the same index, Intel Corporation's (INTC) market capitalization, as of September 2013, is 118,42 billion USD whereas that of Aceto Corporation (ACET) is a mere 416,60 million USD. Hence, investors into stocks of smaller market capitalizations may be less risk-averse, or even be considered more irrational than those seeking safer returns from large-cap stocks.

Table 7
Descriptive Statistics for ISV

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AXE	-0,24	0,00	94,10	-95,34	20,93	0,01	7,33	390,94	0,00
BHI	0,10	0,00	80,44	-73,40	17,53	0,29	6,40	248,19	0,00
BMI	-0,19	0,00	68,06	-61,52	17,47	0,07	4,23	32,21	0,00
BMS	-0,15	0,00	57,18	-60,82	15,55	-0,19	3,92	21,02	0,00
CIR	0,03	0,00	79,69	-77,07	29,99	-0,11	2,94	0,55	0,76
CVX	0,15	0,00	76,41	-71,91	11,20	0,07	15,33	3175,92	0,00
DUK	0,27	0,00	127,30	-123,79	20,49	0,33	10,25	1108,06	0,00
DVN	-0,13	0,00	101,94	-86,11	25,54	-0,12	4,19	31,01	0,00
EMC	0,00	0,00	44,53	-37,11	8,69	0,23	6,93	326,30	0,00
EME	-0,34	0,00	97,27	-86,02	24,81	0,20	4,31	39,23	0,00
F	-0,32	0,00	54,52	-48,34	13,20	-0,05	4,23	32,01	0,00
HAL	-0,13	0,00	104,73	-67,58	15,17	1,12	13,13	2244,52	0,00
HON	-0,18	0,00	25,95	-29,24	7,43	0,09	4,21	31,28	0,00
KO	0,02	0,00	43,43	-35,45	7,40	-0,36	7,96	523,84	0,00
ODP	0,26	-3,87	247,63	-177,43	43,76	0,74	6,45	294,79	0,00
OMX	-0,08	-1,26	34,83	-26,47	8,67	0,53	4,39	63,95	0,00
OXY	-0,21	0,00	30,42	-27,87	7,53	0,32	4,88	82,06	0,00
PFE	-0,27	0,00	86,90	-81,09	14,40	0,30	13,04	2111,52	0,00
PG	-0,15	0,00	54,65	-43,08	13,11	0,09	4,10	26,11	0,00
SO	-0,08	0,00	75,38	-68,06	14,94	0,26	6,23	223,20	0,00
WAG	0,06	0,00	52,61	-42,29	11,39	0,17	5,51	133,97	0,00
WMT	0,27	0,00	53,74	-41,22	10,44	0,39	10,70	1250,51	0,00
AAPL	0,12	0,00	94,16	-69,31	12,65	1,57	15,97	3716,12	0,00
ABFS	0,10	-1,21	70,77	-67,83	20,26	0,13	3,95	18,26	0,00
ACET	0,09	0,00	91,63	-97,65	17,07	-0,02	7,64	449,76	0,00
ACTG	0,12	0,00	32,85	-28,77	8,26	0,09	4,48	46,64	0,00
ADBE	-0,19	0,00	67,69	-78,85	21,00	-0,06	4,11	26,01	0,00
ADI	-0,44	0,00	45,95	-49,72	12,39	0,01	3,70	10,39	0,00
ALTR	-0,12	0,00	51,08	-38,57	10,88	0,17	4,67	60,38	0,00
ASTE	-0,29	0,00	60,61	-76,38	17,39	-0,26	4,76	70,39	0,00
AXAS	-0,26	0,00	59,05	-69,31	15,40	-0,05	5,00	83,77	0,00
BIIB	-0,17	0,00	177,20	-110,87	21,97	1,07	12,80	2102,05	0,00
COST	0,16	0,00	23,44	-27,19	6,96	-0,15	4,95	81,12	0,00
CTXS	-0,18	0,00	51,73	-57,18	9,17	-0,44	14,51	2779,76	0,00
EQIX	-0,09	0,00	74,44	-93,16	24,44	-0,02	4,30	28,95	0,00
EXPE	0,00	0,00	35,20	-15,42	5,91	2,08	11,07	1450,92	0,00
FISV	-0,18	0,00	91,63	-96,28	21,53	0,00	5,38	118,18	0,00
INTC	-0,16	0,00	34,25	-19,85	4,11	1,18	13,83	2562,73	0,00
MYL	0,10	0,00	65,39	-77,65	15,74	0,20	6,13	207,86	0,00
NILE	0,00	0,00	59,47	-78,17	21,30	-0,34	3,86	24,20	0,00
REGN	0,10	0,00	123,97	-94,91	31,29	0,29	3,91	17,36	0,00
SNDK	0,00	0,00	36,10	-32,64	7,87	0,35	7,62	455,45	0,00
VRTX	-0,07	0,00	30,11	-37,22	8,36	-0,20	4,80	0,70	0,00
XLNX	-0,27	0,00	60,22	-47,26	11,36	0,06	7,03	339,47	0,00

	Mean%	Median%	Max%	Min%	SD%	SK	K	J-B	p
AKBNK	0,16	0,00	63,60	-67,25	13,05	-0,08	8,20	574,05	0,00
ALARK	-0,37	0,00	99,68	-98,49	25,79	-0,01	4,93	78,83	0,00
DESA	0,04	0,00	109,86	-82,10	24,57	0,15	3,95	17,31	0,00
PETKM	-0,17	0,00	156,06	-95,10	31,54	0,74	6,10	205,08	0,00
SANKO	-0,10	0,00	96,94	-69,31	23,03	0,37	4,38	42,40	0,00
THYAO	0,06	0,00	97,19	-69,31	18,24	0,24	6,56	273,58	0,00
TOASO	-0,44	0,00	87,55	-78,85	31,14	0,06	2,90	0,44	0,00
YATAS	-0,21	0,00	74,84	-64,80	17,33	0,29	4,78	60,97	0,00

Note: SD, SK, K, J-B and p stand for standart deviation, skewness, kurtosis, Jarque-Bera statistic and its corresponding p-value, respectively.

The average changes in ISV values (standart deviations) for NYSE, NASDAQ and BIST-100 companies are -0,06% (16,35%), -0,08% (14,89%) and -0,13% (23,09%), respectively.

These values imply that ISV data for US stocks has been almost three times more volatile than their corresponding stock returns. This value is even larger for Turkish companies, where ISV data is almost four times more volatile than stock returns. This may suggest in case of a possible effect on stock return volatility that Turkish investor sentiment, on the average, is more volatile than the U.S. investor sentiment.

On a final note regarding descriptives, the three variables, for almost all companies, show serial correlation in their residuals of their respective OLS-regression equations.

4.5. Empirical Findings

This section serves as an illustration for the modelling procedures and analysis rationale applied to all companies covering the U.S. and Turkish, using Intel (INTC) as example and summarizes comprehensive empirical findings.

We start with the pre-testing for ARCH-effects of OLS-Residuals for INTC, which results in an observed r-squared of 28,31 and a corresponding p-value of 0,00. Since this p-value is below our significance level of 0,05, we reject the null hypothesis for this test that there is no ARCH effect and conclude that there is an

ARCH effect in the residuals. A supplemental analysis of normality confirms that the data is not normally distributed with the J-B statistic being 138,79 and corresponding p-value of 0,00. These values imply that since the null hypothesis is that the data is normally distributed, we conclude that it is not normally distributed. Thirdly, we examine the correlogram of squared returns, which shows that serial correlation is present in 35 lags. Hence, we reach the decision that ARCH-type models can be applied.

The pre-testing for ARCH effects, normality and serial correlation in residuals is applied to a total of 88 companies of which 52 pass these tests with respect to the base model and are reported henceforth.

After each model one through six is applied, the diagnostic check of residuals is repeated for each company.

Table 8
Residual Diagnostics for INTC

Pre-Test		**Post-Test Diagnostics**	
ARCH LM			
Obs*R-squared	28,31	Obs*R-squared	7,51
Prob. Chi-Square	0,00	Prob. Chi-Square	0,48
NORMALITY			
Kurtosis	5,57	Kurtosis	4,23
Jarque-Bera	138,78	Jarque-Bera	33,51
p	0,00	p	0,00

Table 8 shows pre- and post-test M1 residual diagnostics for INTC, where the residuals no longer contain ARCH effects after the model is applied. Respective p-values for Ljung-Box Q statistic are all above the set significance level of 5%, confirming that there is no serial correlation left in the residuals. As for normality, while the residual series is still not normally distributed the J-B statistic has reduced significantly.

Models through M1 and M6 and Granger causality analysis are applied in a step-wise manner.

Table 9
Model 1: MM-G(1,1) Results

	ω	*p*	*α*	*p*	*β*	*p*	*AR2*
N-E							
ABFS	0,000	*	0,061	*	0,931	*	0,237
CIR	0,002	*	0,353	*	-0,125	*	0,447
ACTG	0,004		0,014		0,299	*	0,235
CVX	0,000		0,081	*	0,892	*	0,583
REGN	0,003		-0,011		0,270		0,275
VRTX	0,000	*	0,024	*	1,017	*	0,175
DUK	0,000		0,054	*	0,892	*	0,313
EMC	0,000	*	0,126	*	0,774	*	0,351
KO	0,000	*	0,088	*	0,838	*	0,342
EQIX	0,000		0,042		0,878	*	0,438
ADI	0,000		0,030	*	0,965	*	0,449
ALTR	0,000		0,032		0,893	*	0,461
BHI	0,000	*	0,005		0,982	*	0,491
F	0,000	*	0,079	*	0,904	*	0,450
FISV	0,000	*	0,049	*	0,896	*	0,546
BIIB	0,000	*	0,022	*	1,009	*	0,134
SO	0,000	*	0,082	*	0,881	*	0,208
XLNX	0,000		0,017	*	0,980	*	0,450
OXY	0,000		0,088	*	0,849	*	0,542
INTC	0,000		0,045	*	0,793	*	0,542
PFE	0,000	*	0,054	*	0,904	*	0,317
HAL	0,000		0,043	*	0,919	*	0,406
EME	0,000	*	0,022	*	0,962	*	0,539
DVN	0,000		0,053	*	0,917	*	0,401
AXAS	0,000	*	0,088	*	0,855	*	0,211
CTXS	0,000		0,020		0,883	*	0,398
HON	0,000		0,051	*	0,928	*	0,682
MYL	0,000	*	0,100	*	0,860	*	0,301
ASTE	0,001		-0,011		0,765	*	0,372
ACET	0,000		0,024	*	0,958	*	0,205
AXE	0,001	*	-0,046	*	0,942	*	0,489
PG	0,000		0,069	*	0,879	*	0,322
BMS	0,001	*	0,001	*	-0,181		0,441
BMI	0,000	*	0,047	*	0,898	*	0,325

	ω	*p*	*α*	*p*	*β*	*p*	*AR2*
E-B							
AAPL	0,000		0,050	*	0,945	*	0,422
COST	0,000	*	0,071	*	0,882	*	0,331
ADBE	0,000		0,018		0,935	*	0,441
WMT	0,000	*	0,044	*	0,909	*	0,202
ODP	0,000	*	0,066	*	0,917	*	0,420
OMX	0,000	*	0,098	*	0,893	*	0,298
NILE	0,000	*	0,095	*	0,883	*	0,218
SNDK	0,000	*	0,092	*	0,876	*	0,342
EXPE	0,001		-0,024	*	0,781	*	0,290
WAG	0,000		-0,022		0,631	*	0,223

	ω	*p*	*α*	*p*	*β*	*p*	*AR2*
TR							
DESA	0,000	*	0,199	*	0,737	*	0,236
PETKM	0,000	*	0,110		0,602	*	0,344
SANKO	0,001	*	0,411	*	-0,023		0,358
TOASO	0,000	*	0,119	*	0,753	*	0,453
YATAS	0,000	*	0,073	*	0,897	*	0,204
AKBNK	0,000		0,054	*	0,925	*	0,707
ALARK	0,000	*	0,077	*	0,828	*	0,438
THYAO	0,000	*	0,091	*	0,767	*	0,006

Note: "*", "AR2", denote significance "p" at or below 5%, and adjusted r-squared, respectively. "ω","α","β", denote constant, alpha (for ARCH term), and beta (for GARCH term) parameters for the GARCH(1,1) conditional variance equation, respectively. Blank spaces in "p", indicate no significance at or below 5%. Companies are grouped in three groups named N-E (US non-e-businesses), E (US e-businesses) and TR (Turkish companies).

Table 9 depicts the basic market base model with its conditional variance modelled with GARCH (1,1) with neither ISV nor trading volume in the latter equation. All companies have passed diagnostic residual serial correlation and ARCH-LM tests, however, results for only 43 out 52 companies are interpretable due to their violation of positivity constraints imposed upon the GARCH model parameters.

INTC results depict significant ARCH and GARCH parameters with positive values of 0,05 and 0,79, respectively. Meaning that news about volatility and last period's forecast variance, are both contributing to the conditional variance of INTC's stock return. The volatility persistence for the base model sums up to 0,84. This value indicates that the persistence of shocks to the conditional variance persists for a relatively long time and the mean reverts back relatively late to its normal values. This is expected and commonly established, evidenced by the clustering patterns of financial time series. However, INTC's volatility persistence is lower when compared to the group of non-e-businesses holding an average volatility persistence of 0,93. This implies that INTC's reversion to the mean is faster than its comparables group. A sum of both terms lower than or equal to 1 indicates stationarity of the variance. Furthermore, the magnitude of the ARCH parameter is drastically lower than that of the GARCH parameter, emphasizing the autoregressive character of the residual series.

The constant, or the mean return, is 0,01% for INTC and the majority of the stocks, as is expected.

The adjusted r-squared statistic adjusts the coefficient of determination for the number of explanatory terms in the conditional mean. We include this statistic, particularly, to compare the two mean specifications (the basic market model and the AR(1) with the market (index) return as exogenous variable) for their goodness of fit. In INTC's base market model this value is 0,54 compared to the group average of 0,38.

The next step is to analyze the increasingly nested model, which, is Model 2, where we have included the ISV variable to the conditional variance of the previous base model.

With respect to model M2, applied to a total of 52 companies, four companies (SO, ASTE, NILE and DESA) fail to eliminate serial correlation in the residuals post-modelling and ten companies (HAL, REGN, EQIX, ADI, BHI, CTXS, EXPE, WAG, SANKO and ADBE) fail the GARCH parameter positivity constraints. As a result only 38 companies remain for meaningful analysis.

Table 10
Model 2: MM-G(1,1)-ISV Results

	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*AR2*
N-E									
ABFS	0,000	*	0,061	*	0,931	*	0,000		0,238
CIR	0,000	*	0,000		0,926	*	0,002	*	0,444
ACTG	0,003	*	0,023		0,343	*	0,013	*	0,235
CVX	0,000		0,084	*	0,896	*	0,000	*	0,583
REGN	0,006	*	-0,023		-0,482		0,002	*	0,275
VRTX	0,000		0,010	*	0,897	*	0,008	*	0,177
DUK	0,000	*	0,060	*	0,878	*	0,000		0,313
EMC	0,000	*	0,106	*	0,801	*	0,000	*	0,351
KO	0,002		0,075	*	0,868	*	0,000	*	0,342
EQIX	0,002	*	0,170	*	-0,156	*	-0,001	*	0,432
ADI	0,002	*	-0,025	*	-0,798	*	0,000	*	0,451
ALTR	0,000	*	0,027		0,874	*	0,002	*	0,460
BHI	0,020	*	0,008		-0,600	*	0,000	*	0,491
F	0,000	*	0,084	*	0,899	*	0,000	*	0,450
FISV	0,000	*	0,050	*	0,897	*	0,000	*	0,546
BIIB	0,001	*	0,280		0,000	*	0,000	*	0,137
SO	0,000	*	0,080	*	0,880	*	-0,010		0,208
XLNX	0,000		0,017	*	0,979	*	0,000		0,450
OXY	0,000		0,080	*	0,865	*	-0,001	*	0,542
INTC	0,000	*	0,033		0,815	*	0,004	*	0,542
PFE	0,000	*	0,077	*	0,782	*	0,001	*	0,376
HAL	0,002	*	0,115	*	-0,339	*	0,001	*	0,406
EME	0,000	*	0,078	*	0,868	*	0,001	*	0,548
DVN	0,000		0,046	*	0,943	*	0,001	*	0,401
AXAS	0,000	*	0,094	*	0,853	*	-0,007	*	0,238
CTXS	0,003	*	0,045	*	-0,830	*	0,001	*	0,398
HON	0,000	*	0,074	*	0,892	*	0,000	*	0,656
MYL	0,000	*	0,107	*	0,851	*	0,001	*	0,298
ASTE	0,002	*	-0,047	*	0,520	*	0,001		0,372
ACET	0,000		0,016	*	0,968	*	0,004	*	0,200
AXE	0,000	*	0,227	*	0,619	*	0,001	*	0,500
PG	0,000	*	0,068	*	0,874	*	0,000	*	0,323
BMS	0,000	*	0,142	*	0,261		-0,001	*	0,411
BMI	0,000		0,356	*	0,020		-0,001	*	0,279

	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*AR2*
E-B									
AAPL	0,000		0,050	*	0,945	*	0,000		0,422
COST	0,000	*	0,071	*	0,882	*	0,000		0,331
ADBE	0,000		-0,019		0,966	*	0,000		0,441
WMT	0,000	*	0,046	*	0,900	*	0,000		0,202
ODP	0,000	*	0,064	*	0,919	*	-0,001		0,417
OMX	0,000	*	0,099	*	0,892	*	0,000		0,294
NILE	0,007	*	0,054	*	-0,638	*	0,000		0,219
SNDK	0,000	*	0,083	*	0,888	*	0,003	*	0,000
EXPE	0,002		-0,044	*	0,499		0,000		0,288
WAG	0,001	*	-0,066		0,152		0,000		0,223

	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*AR2*
TR			0,115		0,713				
DESA	0,000	*	0,217	*	0,730	*	-0,001	*	0,236
PETKM	0,001	*	0,110		0,521	*	0,001	*	0,344
SANKO	0,001	*	0,408	*	-0,034		0,000		0,358
TOASO	0,000		0,118	*	0,761	*	0,000		0,453
YATAS	0,001	*	0,188	*	0,649	*	-0,003	*	0,205
AKBNK	0,000		0,055	*	0,924	*	0,000		0,707
ALARK	0,000	*	0,075	*	0,835	*	-0,001	*	0,437
THYAO	0,001	*	0,141	*	0,587	*	-0,003	*	0,005

Note: "*", "AR2", denote significance "p" at or below 5%, and adjusted r-squared, respectively. "ω","α","β", denote constant, alpha (for ARCH term), and beta (for GARCH term) parameters for the GARCH(1,1) conditional variance equation, respectively. Blank spaces in "p", indicate no significance at or below 5%. Companies are grouped in three groups named N-E (US non-e-businesses), E (US e-businesses) and TR (Turkish companies).“ψc” denotes ISV parameter.

Table 10 shows INTC's ARCH and GARCH parameters with positive magnitudes of 0,03 and 0,81 respectively. However, as opposed to the base model values, the ARCH parameter has decreased in magnitude from 0,05 and turned insignificant while the GARCH parameter has increased from its previous value of 0,79, respectively. In this particular case, it can be argued that ISV encompasses news about volatility that normally the ARCH term represents while shifting some of its magnitude to the GARCH parameter.

While the volatility persistence for the previous base model INTC was 0,84, through inclusion of ISV in the conditional variance, the volatility persistence has almost stayed constant.

The U.S. non-e-business companies group comparative values for volatility persistence, show a 11% decrease from 0,93 to 0,83 on the average[5]. This implies that inclusion of the investor sentiment causes the mean values of the underlying stock to move relatively faster to their normal levels.

[5] The persistence value of BIIB has been excluded since both GARCH and ARCH effects disappeared through inclusion of ISV rendering volatility persistence as zero and an outlier for average volatility calculation purposes.

Table 11
Model 3: MM-G(1,1)-ISV-Trading Volume Results

	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*ψv*	*p*	*AR2*	*VP M3*	*VP M2*	*VP M1*	*M3-M2*	*M2-M1*
N-E																
ABFS	0,002	*	0,077	*	0,404	*	-0,001		0,004	*	0,240		0,992	0,992		-0,03%
CIR	0,000	*	-0,008		0,929	*	0,000		0,001	*	0,445		0,926			
ACTG	0,003	*	0,031		0,406	*	0,000		0,003	*	0,231	0,406	0,343	0,299	18,39%	14,44%
CVX	0,000	*	0,096	*	0,704	*	0,000	*	0,001	*	0,583	0,800	0,980	0,973	-18,36%	0,72%
REGN	0,001	*	0,040	*	0,721	*	0,000		0,004	*	0,278					
VRTX	0,003	*	0,023	*	0,490	*	0,013	*	0,002	*	0,173	0,513	0,906	1,041	-43,34%	-12,92%
DUK	0,000	*	0,054	*	0,576	*	0,000		0,000	*	0,313	0,630	0,939	0,946	-32,90%	-0,81%
EMC	0,000	*	0,048	*	0,856	*	0,000		0,001	*	0,351		0,908	0,900		0,84%
KO	0,000	*	0,122	*	0,570	*	0,001	*	0,000	*	0,342		0,943	0,925		1,89%
EQIX	0,000	*	0,064	*	0,703	*	0,000		0,001	*	0,433			0,878		
ADI	0,000	*	0,049	*	0,421	*	-0,001	*	0,001	*	0,451			0,995		
ALTR	0,000	*	0,055	*	0,721	*	-0,001	*	0,001	*	0,462	0,776	0,874	0,893	-11,25%	-2,05%
BHI	0,001	*	-0,021	*	0,424	*	-0,001		0,002	*	0,491			0,982		
F	0,001	*	0,146	*	0,562	*	-0,001		0,003	*	0,446	0,708	0,984	0,983	-28,01%	0,06%
FISV	0,000	*	0,072		0,480	*	0,000		0,001	*	0,548		0,947	0,945		0,24%
BIIB	0,000	*	0,042	*	0,852	*	0,001	*	0,002	*	0,136	0,894	0,000	1,031		
SO	0,000	*	0,060	*	0,875	*	0,000	*	0,000	*	0,208			0,964		
XLNX	0,000	*	0,059	*	0,828	*	-0,002	*	0,001	*	0,448	0,887	0,996	0,997	-10,95%	-0,04%
OXY	0,000	*	0,130	*	0,721	*	-0,001		0,001	*	0,542	0,851	0,945	0,937	-9,96%	0,89%
INTC	0,000	*	0,044		0,347		0,003	*	0,001	*	0,533	0,000	0,815	0,838		-2,76%
PFE	0,000	*	0,138	*	0,546	*	0,000	*	0,001	*	0,318		0,859	0,958		-10,29%
HAL	0,001	*	0,111	*	0,319	*	0,000		0,002	*	0,477	0,430		0,962		
EME	0,000	*	0,075	*	0,622	*	0,000		0,001	*	0,536		0,947	0,984		-3,82%
DVN	0,000		0,054		0,899	*	0,000		0,001	*	0,458	0,899	0,989	0,970	-9,04%	1,94%
AXAS	0,002	*	0,085	*	0,540	*	0,002		0,007	*	0,207		0,947	0,944		0,31%
CTXS	0,001	*	0,041		0,293	*	0,002	*	0,001	*	0,388	0,293		0,883		
HON	0,000	*	0,080	*	0,661	*	-0,001	*	0,001	*	0,682		0,966	0,979		-1,37%
MYL	0,000	*	0,153	*	0,725	*	-0,001		0,001	*	0,314	0,878	0,959	0,960	-8,40%	-0,14%
ASTE	0,001	*	0,015		0,568	*	0,002	*	0,002	*	0,370					
ACET	0,001	*	0,088	*	0,386	*	0,001	*	0,001	*	0,205		0,985	0,982		0,31%
AXE	0,001	*	0,002		0,482	*	0,001	*	0,001	*	0,491		0,847			
PG	0,000	*	0,071	*	0,412	*	0,000		0,001	*	0,252	0,483	0,942	0,947	-48,67%	-0,54%
BMS	0,000	*	0,055	*	0,835	*	0,000		0,001	*	0,405	0,890	0,142		527,99%	
BMI	0,002	*	0,152	*	0,091		0,000		0,002	*	0,279	0,152	0,356	0,946	-57,15%	-62,41%
	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*ψv*	*p*	*AR2*	*VP M3*	*VP M2*	*VP M1*	*M3-M2*	*M2-M1*
E-B																
AAPL	0,001	*	0,021	*	0,387	*	0,000		0,003	*	0,339	0,408	0,995	0,995	-59,01%	0,01%
COST	0,000	*	0,161	*	0,359	*	0,001		0,001	*	0,339	0,520	0,953	0,953	-45,43%	-0,01%
ADBE	0,000	*	0,020		0,976	*	0,000		0,001	*	0,434	0,976		0,935		
WMT	0,000	*	0,037		0,444	*	0,001		0,001	*	0,227	0,481	0,946	0,953	-49,14%	-0,68%
ODP	0,002	*	0,170	*	0,392	*	-0,004		0,004	*	0,442		0,983	0,983		0,00%
OMX	0,003	*	0,173	*	0,442	*	-0,010	*	0,006	*	0,304	0,616	0,992	0,991	-37,92%	0,07%
NILE	0,000	*	0,160	*	0,683	*	0,000	*	0,000	*	0,212			0,977		
SNDK	0,002	*	0,000		0,020		0,030	*	0,420	*	0,344		0,970	0,969		0,17%
EXPE	0,002	*	-0,042	*	0,489	*	-0,008		0,004	*	0,290					
WAG	0,000	*	-0,021		0,600	*	0,001		0,001	*	0,223					
	ω	*p*	*α*	*p*	*β*	*p*	*ψc*	*p*	*ψv*	*p*	*AR2*	*VP M3*	*VP M2*	*VP M1*	*M3-M2*	*M2-M1*
TR																
DESA	0,002		0,078		0,480		-0,001		0,001	*	0,237	0,000				
PETKM	0,001	*	-0,032	*	0,493	*	0,001	*	0,001	*	0,344		0,521	0,602		-13,44%
SANKO	0,000	*	0,181	*	0,447	*	0,000		0,000	*	0,355	0,628				
TOASO	0,000		0,114	*	0,753	*	0,000		0,000	*	0,453	0,866	0,879	0,872	-1,39%	0,79%
YATAS	0,002	*	0,140		0,446	*	0,005	*	0,001	*	0,206	0,446	0,837	0,970	-46,68%	-13,70%
AKBNK	0,000		0,121	*	0,542	*	0,000		0,000	*	0,708	0,663	0,979	0,979	-32,29%	0,00%
ALARK	0,001		0,116	*	0,543	*	0,000	*	0,000	*	0,438	0,659	0,911	0,904	-27,64%	0,69%
THYAO	0,002	*	0,080		0,481	*	-0,005	*	0,001	*	0,003	0,481	0,728	0,858	-33,91%	-15,18%

Note: "*", "AR2", denote significance "p" at or below 5%, and adjusted r-squared, respectively. "ω","α","β", denote constant, alpha (for ARCH term), and beta (for GARCH term) parameters for the GARCH(1,1) conditional variance equation, respectively. Blank spaces in "p", indicate no significance at or below 5%. Companies are grouped in three groups named N-E (US non-e-businesses), E (US e-businesses) and TR (Turkish companies).“ψc” denotes ISV parameter.

“ψv” and “VP”denote trading volume parameter and volatility persistence. Blank spaces indicate lack of interpretability due residual serial correlation or failure of non-negativity constraints.

Table 11, shows the results for model M3, which includes both exogenous variables, ISV and trading volume in the conditional variance equation of the error term obtained from the market model. In as such, it is a nested version of the M1 and M2 models. In this model, only 29 out of 52 companies are fit for meaningful analysis since most of the eliminated companies displayed a dramatic increase in serial correlation in their residuals. For the specific groups of the U.S. non-e-businesses while the e-businesses remain for the most part unaffected, and Turkish companies, the former group shows a rough 50% decline in the number of interpretable company results, on the other hand, the percentage for Turkish companies remaining is 70%. Thus, the reduction is not as much pronounced in the Turkish case.

INTC results show that ARCH and GARCH parameters are insignificant. Thus, while inclusion of ISV had rendered the ARCH parameter insignificant, the addition of trading volume has eliminated the GARCH parameter. For the comparable group of companies, the ARCH effect of ISV cannot be generalized, however we can speak of a reduction of the GARCH parameter and a definitely more prounced reduction of such when trading volume is included.

In INTC’s case, the ISV parameter is positively significant when included in isolation to the conditional variance and remains positively significant while slightly decreasing in magnitude when trading volume is included, the latter being significant with a comparatively lower magnitude. As previously mentioned, for almost all companies, when analyzed separately, there is serial dependence in the error terms of trading volume and ISV OLS-residuals. Thus, it is logical that we use similar grounds for arguing that ISV can be fit to explain volatility clustering, just like the information flow proponents do, recalling that these scholars attempt to explain these effects with the serial dependence of the error terms of trading volume.

The adjusted r-squared value for INTC across models one through three is 0,54 on the average and for non-e-business companies it is approximately 0,38.

Next we use the same conditional variance specifications to model the error terms of the AR(1) model with the market (index) return as exogenous variable in the conditional mean: These are models four through six. For brevity purposes, we show a condensed version of the results of all three models in Table 12.

Table 12
Results for Base and Nested AR(1)M-G(1,1) Models

					M6										M5	M4
	ω	***p***	***α***	***p***	***β***	***p***	***ψc***	***p***	***ψv***	***p***	***AR2***	***VP M6***	***ψc***	***p***	***VP M5***	***VP M4***
N-E																
ABFS	0,001	*	0,098	*	0,420	*	-0,001		0,003	*	0,242					0,991
CIR	0,000	*	-0,100		0,935	*	0,000		0,000	*	0,464					
ACTG	0,003	*	0,017		0,370	*	0,000		0,003	*	0,227		0,013	*	0,329	
CVX	0,000	*	0,082	*	0,808	*	0,000	*	0,001	*	0,584	0,890	0,000	*	0,982	0,974
REGN	0,000	*	0,040	*	0,721	*	0,000		0,004	*	0,278		0,001	*		
VRTX	0,003	*	0,027	*	0,494	*	0,014	*	0,001	*	0,176	0,520	0,008	*	0,908	
DUK	0,000	*	0,070	*	0,778	*	0,000	*	0,000	*	0,312	0,849			0,908	0,923
EMC	0,000		0,032	*	0,919	*	-0,001	*	0,001	*	0,353		-0,001	*	0,903	0,902
KO	0,000	*	0,133	*	0,564	*	0,001	*	0,000	*	0,341		0,000	*	0,946	0,926
EQIX	0,000	*	0,036		0,813	*	0,000		0,001	*	0,436	0,813	-0,001	*		0,921
ADI	0,000	*	0,064	*	0,464	*	-0,001	*	0,001	*	0,454		-0,001	*	0,997	0,995
ALTR	0,001		-0,025		0,499		0,002		0,001	*	0,460		0,002	*	0,851	0,886
BHI	0,001	*	0,001		0,090	*	-0,001	*	0,002	*	0,489	0,090	-0,002	*		0,989
F	0,000	*	0,103	*	0,711	*	-0,001		0,002	*	0,453	0,813	0,001	*	0,977	0,974
FISV	0,000	*	0,022		0,448	*	0,000		0,001	*	0,549		0,000	*	0,959	0,951
BIIB	0,002	*	-0,001		0,520	*	-0,002		0,000		0,135	0,520	-0,003	*	0,550	
SO	0,000	*	0,050	*	0,890	*	0,000	*	0,000	*	0,211					0,964
XLNX	0,000	*	0,039	*	0,817	*	-0,001	*	0,001	*	0,453	0,856			0,894	0,928
OXY	0,000	*	0,138	*	0,681	*	-0,001		0,001	*	0,543	0,819	-0,001	*	0,945	0,938
INTC	0,000	*	0,024		0,361		0,003	*	0,001	*	0,531		0,004	*	0,825	0,815
PFE	0,000	*	0,112	*	0,611	*	0,000	*	0,001	*	0,320		0,001	*	0,851	0,956
HAL	0,000	*	0,084	*	0,307	*	0,000		0,002	*	0,479	0,392	0,001	*	0,923	0,959
EME	0,000	*	0,084	*	0,602	*	0,000		0,001	*	0,534		0,000		0,98375	0,98402
DVN	0,000	*	0,057	*	0,895	*	0,000		0,001	*	0,459	0,951	0,001	*	0,984	0,964
AXAS	0,001	*	0,162	*	0,655	*	0,000		0,005	*	0,230		-0,004			0,966
CTXS	0,001	*	0,056		0,320	*	0,002	*	0,001	*	0,386		0,001	*		0,870
HON	0,000	*	0,083	*	0,634	*	-0,001	*	0,001	*	0,683		0,000	*	0,972	0,978
MYL	0,000	*	0,150	*	0,718	*	-0,001		0,001	*	0,304	0,868	0,001	*	0,955	0,957
ASTE	0,000	*	0,042	*	0,743	*	0,001		0,002	*	0,371		0,002	*		
ACET	0,001	*	0,004		0,477	*	0,001		0,001	*	0,213		0,002	*	0,972	0,973
AXE	0,001	*	0,006		0,479	*	0,001	*	0,001	*	0,491		0,002	*	0,624	
PG	0,000	*	0,095	*	0,398	*	0,000		0,001	*	0,257		0,000		0,945	0,957
BMS	0,000	*	0,044	*	0,862	*	0,000	*	0,001	*	0,405	0,907	-0,001	*	0,421	0,953
BMI	0,001	*	0,150	*	0,400	*	0,000		0,002	*	0,280	0,550	0,001	*	0,478	
	ω	***p***	***α***	***p***	***β***	***p***	***ψc***	***p***	***ψv***	***p***	***AR2***	***VP M6***	***ψc***	***p***	***VP M5***	***VP M4***
E-B																
AAPL	0,001	*	0,024	*	0,386	*	0,000		0,003	*	0,342	0,410	0,000		0,993	0,993
COST	0,000	*	0,177	*	0,384	*	0,001		0,001	*	0,341	0,561	0,000		0,966	0,966
ADBE	0,000	*	0,005		0,971	*	0,000		0,001	*	0,436	0,971	0,001			0,951
WMT	0,000	*	0,024		0,461	*	0,000		0,001	*	0,232	0,461	0,000		0,911	0,938
ODP	0,002	*	0,163	*	0,399	*	-0,004		0,004	*	0,444		0,000		0,916	0,983
OMX	0,000	*	0,209	*	0,674	*	0,002		0,002	*	0,291	0,882	-0,001		0,991	0,990
NILE	0,000	*	0,098	*	0,684	*	0,000		0,003	*	0,221		0,000		0,972	0,9737
SNDK	0,002	*	0,104	*	0,431	*	0,001		0,005	*	0,341		0,004	*	0,890	0,990
EXPE	0,002	*	-0,042	*	0,486	*	-0,005		0,004	*	0,287		-0,003			
WAG	0,000	*	-0,017		0,629	*	0,001		0,001	*	0,220		-0,001			
	ω	***p***	***α***	***p***	***β***	***p***	***ψc***	***p***	***ψv***	***p***	***AR2***	***VP M6***	***ψc***	***p***	***VP M5***	***VP M4***
TR																
DESA	0,002	*	0,076		0,479	*	-0,001		0,001	*	0,237		-0,001	*		
PETKM	0,001	*	0,037		0,508	*	0,001	*	0,001	*	0,343	0,508	0,001	*	0,520	0,712
SANKO	0,000	*	0,171	*	0,425	*	0,000		0,000	*	0,358	0,596	0,000			
TOASO	0,001	*	0,173	*	0,411	*	-0,001		0,000		0,450		0,000			0,867
YATAS	0,002	*	0,110	*	0,435	*	-0,005	*	0,001	*	0,220		-0,003	*	0,392	0,972
AKBNK	0,000		0,117	*	0,537	*	0,000		0,000	*	0,708		0,000		0,979	0,979
ALARK	0,000		0,109	*	0,536	*	0,000		0,000	*	0,437		-0,001	*	0,921	0,909
THYAO	0,002		0,086	*	0,495	*	-0,002	*	0,001	*	0,005	0,581	-0,002	*	0,767	0,857

Note: "*", "AR2", denote significance "p" at or below 5%, and adjusted r-squared, respectively. "ω","α","β", denote constant, alpha (for ARCH term), and beta (for GARCH term) parameters for the GARCH(1,1) conditional variance equation, respectively. Blank spaces in "p", indicate no significance at or below 5%. Companies are grouped in three groups named N-E (US non-e-businesses), E (US e-businesses) and TR (Turkish companies)."ψc" denotes ISV parameter. "ψv" and "VP"denote trading volume parameter and volatility persistence. Blank spaces indicate lack of interpretability due residual serial correlation or failure of non-negativity constraints.

Table 12 shows that the adjusted r-squared for all three models four through six, on the average, is 0,38 and for INTC it is approximately 0,53. When compared to the basic market models one through three, these values are almost identical. This implies that both mean specifications are a similarly good fit. This reasoning seems logical since the majority of the expected returns is being explained by the respective market index returns.

While INTC's base model has a significant GARCH parameter of 0,80, it has an insignificant ARCH parameter of 0,04 and no serial correlation left in the post-modelling residuals. With the inclusion of ISV in the conditional variance equation, the ARCH parameter remains insignificant but the GARCH term increases to 0,83 remaining significant, with no serial correlation left in its residuals. The ISV parameter value is a significant 0,004. However, with the inclusion of trading volume, both the ARCH and GARCH terms display no significance with magnitudes of 0,02 and 0,36, while the parameters for ISV and trading volume become 0,003 and 0,001, respectively. What is a crucial result of this model is that, different from the previous, less-nested, two AR(1) models, the residuals of M6 display serial correlation at lags 2-5. Thus, Model 6 is interpreted not to be a good fit in explaining INTC's volatility behavior.

The blank spaces in the VP columns clearly show the reductions in the number of companies in the increasingly nested models. As is in the INTC case, most of this reduction in analyzable companies for Model 3, is due to serial correlation in the residuals that arise when trading volume is added to ISV as a second exogenous variable to the conditional variance.

The percentage reductions of the number of analyzable companies when models with ISV only and ISV and trading volume are compared for Market Model and AR(1)M mean specifications. The reduction percentage for the number of interpretable company results with the market model are mean specifications are

42% for US non-e's, 30% for e-businesses, and 30% for Turkish companies. These same values for the AR(1) market model specification are 48%, 40%, 60% respectively.

While the percentage of US non-e-business, (Turkish) companies with significant ISV parameters in Model M5 are 88% (80%) with 27% (100%) being of positive magnitude, these values are reduced to 43% (67%), respectively, upon inclusion of trading volume. For e-businesses there is no ISV effect in neither the M4 or the M5 model with the exception of one company for the latter.

Thus, in the general case, there is on the average, a 50% reduction of the ISV effect in the conditional variance when trading volume is included for the two national markets, with the U.S. market having a more pronounced effect.

The trading volume parameter, on the other hand, is significant and positive for all 52 companies with the exception of one Turkish company (TOASO) and one US company (BIIB).

In terms of the magnitudes, the average significant net ISV parameter for interpretable US non-e-businesses (and Turkish companies) is 0,002 (-0,002), and, 0,001 (0,000), when trading volume is included in the conditional variance. The average trading volume parameter magnitude as depicted in Model M6 for the U.S. non-e-businesses, e-businesses and Turkish companies is equally 0,001 for all groups. Thus, while there is a decline in companies with significant ISV's for non-e-business and Turkish companies, there is also a change in the magnituded for a part of the remaining firms.

As for volatility persistence for INTC, since we cannot interpret the most nested model, M6, due to serial correlation in residuals, we compare models M4 and M5. In the base model INTC volatility persistence is 0,82, this value becomes 0,83 when ISV is included in the conditional variance equation. However, for the average group of interpretable companies, volatility persistence for non-e-businesses, for models M4, M5 and M6, show a decline and remain approximately; 0,946, 0,843 and

0,703, respectively. That is a percentage decline of volatility of 11% with ISV only in the conditional variance and an additional 17% when trading volume is added.

For Turkish companies, as well, volatility persistence reduces as the model becomes increasingly nested by 19% and 22%, respectively. In contrast, e-businesses show almost no (a mere 2%) reduction in volatility persistence with ISV included, but a drastic 31% with the added trading volume variable.

Thus, the inclusion of trading volume, for the whole group of companies we have analyzed, decreases volatility persistence by more than 30% on the average.

In sum, comparatively across models, with respect to the U.S. non-e-business companies, our empirical findings, for all models, show that volatility persistence displays a decreasing pattern as the model becomes more nested. The inclusion of solely the ISV variable in the conditional variance, leads to a decrease of volatility persistence of approximately 11% and 13%, for market and AR(1)M-type models, respectively. The addition of trading volume as second exogenous variable to the conditional variance enhances this reduction by an additional 17% and 24%, respectively. Thus, both variables cause shocks to volatility to die out quicker and, thus, leading to a faster mean-reversion.

Trading volume, positively affects the conditional volatility of all non-e-business stocks, however it does not eradicate the effect of internet search volume for almost half of the sample.

Similarly, trading volume does not account for G(ARCH) effects, while there is a slight reduction of the magnitudes of the GARCH parameters on the average.

When included in isolation to the conditional variance equation, ISV is significantly affecting almost 90% of the stocks for both mean specifications, and negatively-lenient, in terms of sign, where roughly 30% of the non-e-business group has a positive ISV parameter.

When trading volume is included, the remaining companies show a reduction of the average magnitude of the ISV parameter while, the magnitudes of both variables are similar.

Upon inclusion of trading volume to the conditional variance, the number of analyzable companies declines to nearly half of their original size. The main reason for this reduction, is the striking increase of companies with residual serial correlation, not previously seen in base models or models with solely ISV in the conditional variance. Furthermore, this inclusion results in similar a reduction with almost half of the remaining analyzable companies having a significant ISV effect.

Empirical findings for the set of Turkish companies shows that trading volume, for almost all companies, has a significant additional effect on the conditional variance of the underlying stocks. When trading volume is included into the nested market model, merely three companies lose their significant ISV effect along with GARCH and ARCH effects. When trading volume is included into the nested AR(1)M model, it eradicates the ISV and ARCH term for only one company and ISV, GARCH and ARCH terms for two companies. Again, no general argument that trading volume entirely eradicates effects of ISV, GARCH or ARCH can be made.

Similar to the U.S. non-e business group, inclusion of trading volume as an additional exogenous variable to the conditional variance decreases the number of companies who previously had a significant ISV parameter when analyzed in isolation: Roughly 75% have significant ISV effects with only ISV in the conditional variance. 70% of the companies remain when trading volume is included, with 55% of these having a significant ISV effect.

Overall, as for volatility persistence in the case of Turkish stocks, there is a 13% reduction, on the average, for both mean specifications when ISV is included in the conditional variance, and, another 26%, on the average, comes through the inclusion of trading volume. These values are similar to the group of the U.S. non-e-businesses.

The control group of e-businesses displays dramatically different results, with respect to ISV effect, relative to the US non-e-businesses. As such, with the exception of one, none displays any ISV effect. Along these lines, the volatility persistence is also almost unaffected by inclusion of ISV to the base models.

Similar to the other two groups, trading volume, for the whole group of e-businesses as well, has a positive and significant effect on conditional variance.

Prior to causality testing, we need to determine the lag specification. This is performed through a VAR analysis of related variables for each and every one of the 52 companies separately. Afterwards, the lag structure is examined and appropriate lag length chosen according to the minimum AIC value.

Table 13
VAR Lag Order Selection for INTC

Lag	AIC	
0	-4,2196	
1	-4,2333	*
2	-4,2304	
3	-4,2268	
4	-4,2262	
5	-4,2243	
6	-4,2266	
7	-4,2235	
8	-4,2201	

Note: "*" denotes 5% significance

Table 13 depicts the lowest AIC and specifies the lag for INTC. In INTC's case it is lag 1. The AIC is chosen as a criterion for lag order selection as opposed to SIC relying on Ivanov and Lilian (2005) judgements that it produces the most realistic results with sample sizes like ours. Lag specifications for the remaining companies is displayed in the right-most column of Table 14.

Table 14
Granger Causality Test Results

N-E	P--> ISV	ISV--> P	ISV--> V	V--> ISV	Lag
ABFS		*	*		1
CIR				*	2
ACTG			*		2
CVX			*	*	7
REGN					3
VRTX	*			*	2
DUK		*		*	2
EMC			*	*	2
KO			*	*	2
EQIX				*	2
ADI				*	6
ALTR				*	3
BHI				*	2
F					2
FISV				*	2
BIIB	*		*	*	1
SO					2
XLNX			*	*	1
OXY				*	2
INTC					1
PFE			*	*	2
HAL					2
EME				*	2
DVN	*	*		*	1
AXAS				*	2
CTXS			*	*	2
HON				*	3
MYL					8
ASTE			*	*	1
ACET					2
AXE				*	2
PG			*	*	6
BMS				*	2
BMI					2
E-B	**P--> ISV**	**ISV--> P**	**ISV--> V**	**V--> ISV**	**Lag**
AAPL				*	2
COST				*	2
ADBE					2
WMT				*	1
ODP					5
OMX					6
NILE					2
SNDK	*		*	*	8
EXPE					2
WAG					6
TR	**P--> ISV**	**ISV--> P**	**ISV--> V**	**V--> ISV**	**Lag**
DESA					1
PETKM					2
SANKO			*	*	2
TOASO					5
YATAS				*	4
AKBNK					2
ALARK					1
THYAO					1

Note: "P","ISV","V" and "lag" denote stock price return, change in ISV and trading volume, respectively. “*” denotes 5% significance.

The results of Granger causality analysis with the respective lag order is depicted in Table 14.

Accordingly, INTC shows no temporally causal relationship, results for the rest of the companies are mixed. While for 70% of the US non-e-business companies changes in trading volume precede changes in ISV (at varying lags), for 33% of this same group the temporal ordering is vice versa. For 27%, on the other hand, there is a bi-directional relationship between ISV and trading volume. With the exception of one company (DVN), there is no bi-directional between price and ISV. Roughly 9% of this group shows a uni-directional Granger causal relationship from price to ISV and from ISV to price.

For the e-business group, only one company (SNDK) shows a uni-directional Granger causal relationship from price to ISV and a bi-directional relationship between ISV and trading volume. However, since its lowest AIC is determined to be at lag 8, Granger causality results for SNDK, too, pertain to lag 8 and are, thus, not clearly interpretable.

In the case of Turkish companies, only SANKO displays a bi-directional temporal ordering of ISV and trading volume at lag 2, while trading volume changes precede ISV changes for YATAS at lag 4.

With regard to Granger causality test results for e-businesses, 40% of the sample displays a uni-directional relationship with changes in trading volume preceding changes in ISV.

At the outset of this study, we specified two alternative conditional mean equations. The parameters of the autoregressive term along with the parameters of market (index) return are reported in Table 15.

Table 15
Market and AR(1) Parameters for Model M6

	λ	*p*	δ	*p*		λ	*p*	δ	*p*		λ	*p*	δ	*p*
NASDAQ					**NYSE**					**TR**				
AAPL	1,09	*	-0,04		AXE	1,33	*	0,02		AKBNK	1,24	*	-0,05	
ABFS	1,24	*	-0,05		BHI	1,40	*	-0,08		ALARK	0,75	*	-0,01	
ACET	0,86	*	-0,09	*	BMI	1,35	*	-0,08		DESA	0,79	*	-0,05	
ACTG	1,31	*	-0,06		BMS	0,82	*	-0,03		PETKM	0,68	*	-0,02	
ADBE	1,04	*	-0,05		CIR	1,55	*	0,04		SANKO	0,64	*	-0,07	
ADI	0,97	*	-0,08		CVX	0,95	*	0,05		THYAO	0,15	*	-0,02	
ALTR	1,12	*	0,00		DUK	0,47	*	0,05		TOASO	1,08	*	-0,09	
ASTE	1,48	*	0,08		DVN	1,19	*	-0,04		YATAS	0,67	*	0,11	
AXAS	1,49	*	0,01		EMC	1,00	*	-0,09	*					
BIIB	0,67	*	-0,04		EME	1,02	*	-0,05						
COST	0,60	*	-0,03		F	1,57	*	-0,08						
CTXS	1,14	*	0,00		HAL	1,35	*	-1,10	*					
EQIX	1,24	*	-0,12	*	HON	1,15	*	-0,04						
EXPE	1,16	*	0,02		KO	0,52	*	-0,06						
FISV	0,89	*	-0,10		ODP	2,31	*	-0,02						
INTC	1,00	*	-0,07		OMX	1,45	*	0,02						
MYL	0,80	*	0,01		OXY	1,25	*	-0,07						
NILE	0,17	*	-0,02		PFE	0,68	*	-0,04						
REGN	1,28	*	0,08		PG	0,38	*	-0,04						
SNDK	1,80	*	0,05		SO	0,35	*	-0,06						
VRTX	1,04	*	0,05		WAG	0,63	*	-0,02						
XLNX	1,03	*	-0,05		WMT	0,45	*	-0,05						

Note: λ and δ are parameters for market and AR(1) terms, respectively. "*" denotes at or below 5% significance.

As shown in Table 15, the market model parameter average magnitudes are 1,06 and 1,05 respectively for NASDAQ and NYSE-companies. The Turkish one is lesser in magnitude and approximates to an average value of 0,75. The autoregressive components, however, are only significant for two companies in each US market.

Accordingly, across all companies the previous lag of the stock return are, in the majority of cases, an insignificant explanatory variable in the conditional mean equation.

Recalling, from our previously reported results, that models with two alternative mean specifications display similar effects of ISV and trading volume on the conditional variance equation, one may infer that the autoregressive component, seems not make a major difference in terms of effect on linear mean returns when included along with the market returns in the conditional mean equation.

Empirical evidence seems to mainly support our research hypotheses. Particularly, research hypothesis H_1, which states that there is a significant ISV impact on conditional volatility, is supported in the case of the U.S. non-e-businesses and Turkish companies. For e-businesses, this research hypothesis, as expected, can be rejected.

The research hypothesis H_2, stating that ISV and trading volume significantly affect conditional variance, is supported as well, for both, the U.S. non-e-businesses and Turkish companies. For e-businesses, it is partially supported, in that only trading volume affects conditional variance, with ISV having no noticeable effect on such, as foreseen.

As for H_3, there is no dominant evidence of a temporal ordering with respect to the Turkish and e-business groups.

H_4 is supported for, both, the U.S. non-e-businesses and Turkish companies, as expected.

For e-businesses, as expected, volatility persistence decreases only due to the inclusion of trading volume with ISV having virtually no effect. This latter manifestation, is another support for the argument of the adequacy of using e-businesses as control groups.

This leads us to the final research hypothesis, H_5, which states that e-businesses can be used a control groups to justify usage of name-based queries. Since all indicators speak for the lack of effect of the ISV variable on the conditional variance with regard to e-businesses, H_5 is supported.

CHAPTER 5
CONCLUSION

The beginning of this paper presented a detailed discussion surrounded by the arguments of the proponents of classical finance, arguing that markets are efficient and there are no discernible patterns, and thus, no predictability in price movements of equities.

However, as empirical findings suggesting the contrary emerged over time, classical finance suffered attacks coming especially from the behavioral finance front. Scholars started embracing the idea that the decision making process of investors may not necessarily be fully rational and that individuals were limited in their capabilities of acquiring, processing and acting upon information which led to the popularization of concepts known as heuristics, or simply, mental shortcuts.

As such, behavioral finance literature embarked on a quest for "clues" that might reveal what investors were thinking. While some scholars thought of these investors as irrational, others considered them rationally-bounded. Yet, another strand of literature argued that, these were rational investors engaging in tactical investing based on the anticipated behavioral patterns of irrational investors. Thus, there is no common agreement as to who it is that causes prices to deviate from their fundamentals and exhibit volatility clustering behavior. In that realm, we have expanded upon the obscure concept of "noise", open to debate, both in the statistical sense be it as error term, residual, or idiosyncratic risk, and, its behavioral meaning, used frequently to delineate a signal other than information. An appendage to the concept of noise, the noise trader theories, were touched upon to complement the theoretical framework of this study.

Next, we presented the issue of the intermittent search for proxies of investor sentiment, a term used to describe the fact that the limited or irrational component of the human psyche is at play when investing, as it is in the entire day-to-day activities of individuals. Over the course of time many such proxies have emerged, we have called them "traditional" throughout the text. One of the most often studied variables has been stock trading volume, predominantly considered as a proxy of information flows, frequently associated with the mixture of distributions hypothesis.

To contribute to volatility studies, we sought to present a novel proxy of investor sentiment, internet search volume, which was first discovered in academic research as recent as only a few years ago. Still, there is a lot of room for research in behavioral finance in general, and, inclusion of this novel proxy in conjunction with traditional ones, in particular.

To refrain from a U.S.-centric perspective, we included data on the Turkish market as well, despite its limitations. Thus, we were able to present some comparisons between developed and emerging markets.

To that end, fuiled by the awareness of the heterogenous character of our units of analysis, we used an unprecedented census group. We started with a universe of more than 3000 companies belonging to three different indices, and were left with a group of 52 firms, to portray the most reliable and objective results possible.

5.1. Discussion of Findings

In light of our analysis, we have found strong evidence, that investor sentiment, as proxied by ISV does affect stock return volatility. This argument applies especially to the U.S. stocks and to a relatively slightly lesser, though still discernible degree, to the Turkish stocks.

One reason for this latter determination might be due to the data source; Google Trends. For instance, a U.S. company may most certainly receive relatively more hits than a comparable Turkish company. Thus, although they may have a similar query index values this does not indicate that they receive “hits” by the same number of internet users. Along similar lines, the composition of investors and internet users for either company may differ a lot, such that, it may be possible that more traditional investors acting as noise traders but not using the internet may be investing in Turkish companies.

Another argument can be that our sample of Turkish companies could also be divided into e-businesses and non-ebusinesses. For instance Akbank, which is one of the largest commercial Turkish banks, can be viewed as an e-business, serving as a potential explanation for its lack of the investor sentiment effect.

While literature predominantly seems to be in agreement that trading volume has a significant effect on stock return volatility, when analyzed together with ISV, we have found that the significance of the latter seems to remain in tact for half of the sample. Also the magnitudes of the ISV and trading volume parameters are similar. This implies that ISV, in addition to trading volume serves as a significant variable to explain conditional volatility and should be included in future volatility studies.

This study substantiates findings, with respect to the U.S. market, that trading volume decreases volatility persistence, and, to a lesser extent, the magnitude of GARCH parameters. This is especially true for the U.S. market, represented by the U.S. non-e-business and e-business groups.

For Turkish companies, the inclusion of trading volume and its effect on the decrease of volatility persistence, is still large, although the reduction in the number of companies due to serial correlation problems, is not as pronounced as in the U.S. case.

Thus, information flow and internet search behavior seem to be likely related to some extent, more in the U.S. case. This may be due to a combination of relative differences between demographic characteristics of online users, noise traders, percentage of investors with access to the internet, risk-taking behavior and the different levels of informational efficiency of the two markets.

Since there is no apparent temporal ordering as a result Granger causality analysis, for the majority of the underlying stocks, we cannot speak of a significant bilateral interaction reflecting, for instance, any return chasing behavior. Exception to this rule are three companies, where changes in internet search behavior precede changes in price at lags two and one.

This low percentage, we can argue, may be substantiated by alternative mean-specification analyses, showing approximately 20% of the sample to have a linear relationship between price and ISV changes.

The only predominant finding of temporal ordering is the uni-directional causality between ISV and trading volume of U.S. non-e-business companies, found for almost 70% of this group. On the other hand, for more than 30% of this group there is also a uni-directional relationship from ISV to trading volume. But, as explained previously, we cannot make a generalization regarding causality since these rely on extended lags and are very difficult to interpret in light of the fact that we have used data at weekly frequency.

However, this finding when taken together with the results of the GARCH analyses, lead us to reason that these two variables may be interacting to a certain degree. For instance, investors performing the searches may either themselves engage in trade, or, cause the searched keyword to move higher up in the visibility rankings, thereby creating a framing bias for other investors who likewise go on to trade on that information, both, or either, affecting trading volumes. Changes in trading volume on the other hand, may catch people's attention who go online to search for that specific company and act upon their confirmation bias when the name of the most searched company, suggested by the autocomplete function of Google, turns out to be their object of interest. It could also be argued that these two variables are partially driven by a latent variable, for instance, the supply of information provided by the internet or other media channels.

E-businesses, our control group, as opposed to the other two groups, display no noticeable ISV effect on conditional variance, while, as expected, the inclusion of trading volume dramatically reduces volatility persistence.

Apart from statistical underpinnings for using two alternative mean specifications, the market model was chosen due to its popular usage, whereas the second one with the added autoregressive term we consider to be theoretically more adequate for behavioral finance studies. As a result, neither model emerges better than the other, and, results of our analyses do not change much for either model. One likely reason is that most of the expected stock returns are accounted for by the exogenous market (index) return variable, which is included in the conditional means of both models.

Lastly, there are no apparent differences pertaining to companies belonging to either the NASDAQ- or NYSE composite indices, in terms of ISV and trading volume effects on their average volatility behavior. For instance looking at the most nested models both the NASDAQ- and NYSE-composite stocks display an average volatility persistence of 0,65. For the lesser nested models this value, on the average, 0,86.

5.2. Limitations

As with all academic research we are limited by the availability and format of the data, which volatility analysis is very sensitive to. ISV data provided by Google Trends is available only in weekly frequency. This limitation is particularly evident when volume-return volatility studies are considered, which commonly use daily and even intra-daily data.

Furthermore, at this point in time, Google Trends data is not offered for pre-2004 periods, which would have been good to investigate, especially within realm of the Internet Bubble of the early 2000s. However, this latter limitation can be remedied by reasoning that since information seekers have been using the internet actively only for the past decade; internet search volume would not have been a good proxy. Thus, the time period of our analysis is probably the most adequate for measuring investor sentiment through internet search volume.

Internet search volume data for Turkish companies is clearly not as comprehensive, widespread and readily available as for the U.S. companies. As this thesis is concluded, a new update of Google Trends is already in place. Since technology is evolving exponentially these days, we may not be too far away from more reliable and available data on emerging markets.

Another evident limitation discussed in the data and sampling section, is that most stocks have generic names and are thus, naturally omitted from the study.

5.3. Contribution

This study belongs to a newly emerging group of behavioral finance literature focusing on models that integrate ISV as investor sentiment variable. The contributions can be summarized as follows:

First, the use of name-based ISV data, and, its justification through the e-business control group, is unprecedented. A very interesting manifestation and novel contribution to the few papers on internet search volume, is the usage of e-businesses as control groups to show that name-based search queries do represent individual investors as opposed to consumers. The existing few previous papers we have mentioned in Section 2.6 of this study, use either ticker symbols like Da, Engelberg and Gao (2011) or names as in the case of Bank, Larch and Peter (2011), but, to the best of our knowledge, do not provide empirical justification of such.

When a distinction is drawn between companies, whose revenue is obtained mostly from online transactions and those, who operate through physical stores, the analysis results confirm the rationale of arguing that name-based search queries are very likely to represent the individual, irrational or rationally-bounded, investor. As such our main contribution is to bring literature a step closer to the debated issue of the Noise Trader Model by Delong et al. (1990), and the identification of the noise trader as individual investor. In that regard we agree with Da, Engelberg and Gao (2011), that ISV can be used as a proxy for noise trader sentiment of individual investors.

Along the lines of the study by Latoeiro, Ramos and Veiga (2013), we conclude that search behavior is likely to be associated with an action and provide support for Barber, Odean and Zhu (2006), who claim that aggregate individual investor trading must be systematic.

Second, to the best of our knowledge, it is the first study using ISV and a traditional investor sentiment proxy with alternative mean specifications applied to two different national markets over the broadest time period possible. The comparative analysis between a developed and an emerging market provides novel insights, both to the trading volume-variance literature, and, to the ISV literature. Partially comparable previous studies on grounds of the time period they use are those of Da, Engelberg and Gao (2011), Bank, Larch and Peter (2011) and Latoerio, Ramos and Vega (2013) which apply VAR and regression metholodology, respectively. However, the time period they analyze are still very narrow compared to the present study, and the stock price data they peruse belongs to developed markets only.

Third, we argue and demonstrate, that usage of world-wide company name search queries provides better data for analysis in the case of the U.S. companies while the opposite is true for Turkish companies. This is an insightful determination for future studies.

Fourth, we substantiate prior arguments that, stocks belonging to different markets may vary in their conditional variance behavior. In this regard, we are inclined to agree with the stream of literature arguing that emerging markets display different volatility patterns than the more developed markets.

The present study is also mainly consistent with the findings of Baklacı et al. (2011) of a significant and positive trading volume on GARCH effects. The inclusion of the news dummy, as suggested by the authors, as well as an interaction variable of such with trading volume and ISV, might provide further valuable contributions.

Fifth, we confirm findings of the literature following Lamoureux and Lastrapes (1990), Brailsford (1996), and, Omran and McKenzie (2000), that inclusion of trading volume decreases volatility persistence of underlying stocks.

We disagree with Lamoureux and Lastrapes (1990) that trading volume, when included in the conditional variance, eradicates G(ARCH) effects and concur with Omran and McKenzie (2000).

This is overtly visible in the e-business group sample, where the ISV has no effect on the conditional variance.

Also, an important contribution of the present study to the MDH-volume literature is that trading volume and ISV together and significantly also exert a significant effect upon the conditional variance of the underlying stocks and lower the volatility persistence. It also deems important from a noise trader perspective to emphasize that, trading volume data are ex-post outcomes encompassing trades executed by all types of traders, be it individuals or institutions, rational, irrational, or rationally-bounded. ISV data in contrast, is highly likely to represent the individual investor.

Sixth, we show that there is likely to be some level of interaction between internet search behavior and information flow and there is some bi-directional causal relationship, whose nature, however, needs to be explored in other studies which is beyond the scope of this thesis. This argument is supported by findings of Da, Engelberg and Gao (2011), who determine that trading volume is related to ISV but explains only a small part of its variation.

Seventh, we apply an original and very thorough selection process for the companies of analysis, requiring large amounts of labour and time. At the outset the whole population of NASDAQ- and NYSE-composite, and BIST-100 companies, is taken and subjected to only one subjective criteria (the eyeball test). The rest of the criteria were basically based on data availability. In that regard we have analyzed a census rather than a sample and argue that our results are quite reliable.

5.4. Suggestions for Further Research

The findings and methodology of the present study can be used as a foundation for further fruitful research in distinct areas.

Using the same dataset the methodology can be enriched and findings of other models be compared with the present ones. Most of the time series data, originally of non-stationary character, is subsequently logarithmically transformed to become stationary. Further studies may, for instance, use cointegration tests and avoid losing data points through transformation.

Also, there are many companies cross-listed on various exchanges, such as EBAY stock, being traded on various exchanges like Hamburg, Mexico and Sao Paolo. All these price returns from the same stock may be used as dependent variable in a multivariate GARCH framework and analyzed by including both, world-wide and regional ISV data.

In this study, ISV is obtained through Google Trends, since it has the largest market share of the search engine market. However, there are also search engines like Yahoo, Bing and the Russian Search Engine, Yandex. Data from such, once provided in analyzable format, could be used in combination with the Google-provided ISV to account for the sentiment of the whole search engine market.

As illustrated in the present study, we empirically investigate our proposed novel proxy, ISV, together with a traditional proxy. There are obviously numerous alternative variables which can be tried in addition to these.

In the realm of the present study, there is no e-business control group for Turkish companies, since most e-businesses in Turkey are not publically traded yet. Once a sizeable number of Turkish e-business data becomes available, we suggest filling the gap of our data-imposed limitation.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, *19*(6), 716-723.

Alexander, S. S. (1961). Price movements in speculative markets: Trends or random walks. *Industrial Management Review*, *2*, 7-26.

Amsler, C. (1984). Term structure variance bounds and time varying liquidity premia. *Economics Letters*, *16*(1), 137-144.

Antón, M., & Polk, C. (2010). Connected Stocks. *London School of Economics Working Paper*. *FMG Discussion Paper 651*.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, *59*(3), 1259-1294.

Arbel, A. (1985). Generic stocks: An old product in a new package. *The Journal of Portfolio Management*, *11*(4), 4-13.

Askitas, N., & Zimmermann, K. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly, Duncker & Humblot, Berlin*, *55*(2), 107-120.

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance*, *55*, 2219-2257.

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, *7*(3), 271-299.

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*(4), 1645-1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21 (2007), 129-152.

Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, *104*(2), 272-287.

Baklaci, H. F., Tunc, G., Aydogan, B., & Vardar, G. (2011). The Impact of Firm-Specific Public News on Intraday Market Dynamics: Evidence from the Turkish Stock Market. *Emerging Markets Finance and Trade*, *47*(6), 99-119.

Bali, T. G., & Cakici, N. (2008). Idiosyncratic volatility and the cross section of expected returns. *Journal of Financial and Quantitative Analysis*, *43*(1), 29.

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial markets and portfolio management*, *25*(3), 239-264.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261-292.

Barber, B., Odean, T., & Zhu, N. (2006, September). Do noise traders move markets?. In *EFA 2006 Zurich Meetings Paper*.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, *21*(2), 785-818.

Barber, B. M., Odean, T., & Zhu, N. (2009). Do retail trades move markets?.*Review of Financial Studies*, *22*(1), 151-186.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053-1128.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307-343.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance*, *32*(3), 663-682.

Beer, F., Watfa, M., & Zouaoui, M. (2011). *Do investors care about noise trader risk?* (No. 1111201). Université de Bourgogne-Leg (laboratoire d'économie et de gestion)/Fargo (Research center in Finance, organizational Architecture and Governance).

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, *13*(1), 1-42.

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The quarterly journal of Economics*, *110*(1), 73-92.

Bennett, J. A., Sias, R. W., & Starks, L. T. (2003). Greener pastures and the impact of dynamic institutional preferences. *Review of Financial Studies*,*16*(4), 1203-1238.

Bernoulli, Daniel, 1738. Specimen theoriae novae de mensura sortis. Comentarii Academiae Scientiarum Imperialis Petropolitanae, 5, 175–192.

Berrada, T., & Hugonnier, J. (2012). Incomplete information, idiosyncratic volatility and stock returns. *Journal of Banking & Finance*.

Black, F. (1976). The pricing of commodity contracts. *Journal of financial economics*, *3*(1), 167-179.

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 529-543.

Black, F. (1993). Estimating expected return. *Financial Analysts Journal*, 36-38.

Black, F. (1993). Return and beta. *Journal of Portfolio Management*, *20*, 8-8.

Bloomfield, R., O'Hara, M., & Saar, G. (2009). How noise trading affects markets: An experimental analysis. *Review of Financial Studies*, *22*(6), 2275-2302.

Bohl, M. T., Brzeszczyński, J., & Wilfling, B. (2009). Institutional investors and stock returns volatility: Empirical evidence from a natural experiment. *Journal of Financial Stability*, *5*(2), 170-182.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1-8.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, *11*(2), 143-172.

Bollerslev, T., & Jubinski, D. (1999). Equity trading volume and volatility: latent information arrivals and common long-run dependence. *Journal of Business and Economic Statistics* 17,9–21.

Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., & Weber, I. (2012). Web search queries can predict stock market volumes. *PloS one*, *7*(7), e40014.

Brailsford, T. J. (1996). The empirical relationship between trading volume, returns and volatility. *Accounting & Finance, 36*(1), 89-111.

Brandt, M. W., Brav, A., Graham, J. R., & Kumar, A. (2010). The idiosyncratic volatility puzzle: Time trend or speculative episodes?. *Review of Financial Studies*, *23*(2), 863-899.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of financial Economics*,*7*(3), 265-296.

Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge University Press.

Brooks, R. D., Faff, R. W., & Fry, T. R. (2001). GARCH modelling of individual stock data: the impact of censoring, firm size and trading volume. *Journal of International Financial Markets, Institutions and Money, 11*(2), 215-222.

Brown, G. W. (1999). Volatility, sentiment, and noise traders. *Financial Analysts Journal*, 82-90.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, *11*(1), 1-27.

Brown, G. W., & Cliff, M. T. (2005). Investor Sentiment and Asset Valuation*.*The Journal of Business*, *78*(2), 405-440.

Campbell, J. Y., & Kyle, A. S. (1993). Smart money, noise trading and stock price behaviour. *The Review of Economic Studies*, *60*(1), 1-34.

Campbell, J. Y., & Shiller, R. J. (1988). Interpreting cointegrated models. *Journal of Economic Dynamics and Control*, *12*(2), 505-522.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance*, *43*(3), 661-676.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics*,*108*(4), 905-939.

Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance, 56*(1), 1-43.

Canarella, G., & Pollard, S. K. (2011). The Mixture Of Distribution Hypothesis And The Russian Stock Market. *Journal of Business & Economics Research (JBER)*, *1*(11).

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, *52*(1), 57-82.

Chaudhuri, K., & Koo, K. (2001). Volatility of Stock Returns: Importance of Economic Fundamentals. *Economic and Political Weekly,* 3852-3856.

Chemmanur, T. (2009). *Advertising, attention, and stock returns* (Doctoral dissertation, School of Business Administration, Fordham University).

Chichernea, D., Petkevich, A., & Reca, B. (2012). Idiosyncratic Volatility, Institutional Ownership, and Investment Horizon. *Institutional Ownership, and Investment Horizon (October 12, 2012)*.

Choi, H., & Varian, H. (2009). Predicting initial claims for unemployment benefits. *Google Technical Report. Retrieved June 2013, http://static.googleusercontent.com/media/research.google.com/tr//archive/papers/initialclaimsUS.pdf*

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record, 88*(s1), 2-9.

Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41, 135–156.

Cohen, R. B., Gompers, P. A., & Vuolteenaho, T. (2002). Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics*, *66*(2), 409-462.

Comin, D. A., & Philippon, T. (2006). The rise in firm-level volatility: Causes and consequences. In *NBER Macroeconomics Annual 2005,20*, 167-228. MIT Press.

Cooper, C. P., Mallon, K. P., Leadbetter, S., Pollack, L. A., & Peipins, L. A. (2005). Cancer Internet search activity on a major search engine, United States 2001-2003. *Journal of Medical Internet research*, *7*(3).

Coval, J. D., & Shumway, T. (2001). Expected option returns. *The Journal of Finance*, *56*(3), 983-1009.

Da, Z., Engelberg, J., & Gao, P. (2010). The sum of all fears: investor sentiment and asset prices. *SSRN eLibrary*.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461-1499.

D'Amuri, F., & Marcucci, J. (2010). *'Google it!'Forecasting the US unemployment rate with a Google job search index* (No. 2010, 31). Nota di lavoro//Fondazione Eni Enrico Mattei: Global challenges.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance, 56*(3), 921-965.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, *52*(3), 1035-1058.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance, 53*(6), 1839-1885.

DeBondt, W. 8: Thaler, R.(1985). Does the stock market overreact. *Journal of Finance*, *42*, 557-581.

DeLong, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1987). The economic consequences of noise traders. *NBER Working Paper*, 2395.

DeLong, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379-395.

DeLong, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1991). The survival of noise traders in financial markets. *Journal of Business, 64* (1), 1-19.

Depken, C. A. (2001). Good news. bad news and GARCH effects in stock return data. *Journal of Applied Economics*, *4*(2), 313-327.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Dimpfl, T., & Jank, S. (2011). *Can internet search queries help to predict stock market volatility?* (No. 11-15). *CFR Working Paper.*

Eiling, E. (2006). Can nontradable assets explain the apparent premium for idiosyncratic risk? The case of industry-specific human capital. *Unpublished Working Paper. Tilburg University, Netherlands*.

Elton, E. J., & Gruber, M. J. (1977). Risk reduction and portfolio size: An analytical solution. *The Journal of Business, 50*(4), 415-437.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.

Engle, R. F., & Patton, A. J. (2001). What good is a volatility model. *Quantitative finance*, *1*(2), 237-245.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica: Journal of the Econometric Society*, 391-407.

Epps, T.W. & Epps, M.L. (1976). The stochastic dependence of security price changes and transaction volumes: implications for the mixture of distributions hypothesis. *Econometrica,* 44(2), 305-321

Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. *The Journal of Business*, *36*(4), 420-429.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34-105.

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

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, *49*(3), 283-306.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Fama, E. F., & French, K. R. (2002). The equity premium. *The Journal of Finance*, *57*(2), 637-659.

Fama, E., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review, 10*.

Feller, W. (1950). Probability theory and its applications. *New York*, 118-123.

Festinger, L., Riecken, HW, & Schachter, S. (1956). *When prophecy fails*. Minneapolis: University of Minnesota Press

Fink, J., Fink, K., Grullon, G., & Weston, J. P. (2005). Firm age and fluctuations in idiosyncratic risk. *Working Paper, James Madison University.*

Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies*, *15*(3), 751-782.

Fleming, J., Kirby, C., & Ostdiek, B. (2004). Stochastic volatility, trading volume, and the daily flow of information. *Rice University, Jones Graduate School Working Paper*.

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual investors and volatility. *The Journal of Finance*, *66*(4), 1369-1406.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of Financial Economics*, *8*(1), 55-69.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, *19*(1), 3-29.

Friedman, M. (1953). *Essays in positive economics* (Vol. 231). University of Chicago Press.

Gallo, G. M., & Pacini, B. (2000). The effects of trading activity on market volatility. *The European Journal of Finance*, *6*(2), 163-175.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *Review of Financial Studies*, *14*(1), 1-27.

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The High‐Volume Return Premium. *The Journal of Finance*, *56*(3), 877-919.

Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance*, *38*(1), 1-33.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2008). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012-1014.

Girard, E. & Biswas, R. (2007). Trading volume and market volatility: developed versus emerging stock markets. *The Financial Review*, 42, 429-459.

Girma, P.,& Mougoue, M. (2002). An empirical examination of the relation between futures spreads volatility, volume, and open interest. *Journal of Futures Markets* 22, 1083–1102.

Givoly, D., & Palmon, D. (1985). Insider trading and the exploitation of inside information: Some empirical evidence. *Journal of Business*, 69-87.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779-1801.

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., & Watts, D. J. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences*, *107*(41), 17486-17490.

Goyal, A., & Santa‐Clara, P. (2003). Idiosyncratic risk matters!. *The Journal of Finance*, *58*(3), 975-1008.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37 (3): 424–438.

Greenwood, R., & Thesmar, D. (2011). Stock price fragility. *Journal of Financial Economics*, *102*(3), 471-490.

Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization*, *17*(1), 31-57.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393-408.

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *Review of Financial Studies*, *17*(2), 439-461.

Gursoy, G., Yuksel, A., & Yuksel, A. (2008). Trading volume and stock market volatility: evidence from emerging stock markets. *Investment Management and Financial Innovations, 5*(4), 200-210.

March, J. G., & Simon, H. A. (1958). Organizations. Wiley.

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, *58*(3), 1009-1032.

Ho, C., & Hung, C. H. (2009). Investor sentiment as conditioning information in asset pricing. *Journal of Banking & Finance*, *33*(5), 892-903.

Hou, K., Xiong, W., & Peng, L. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum.

Ivanov, V., & Kilian, L. (2005). A practitioner's guide to lag order selection for VAR impulse response analysis. *Studies in Nonlinear Dynamics & Econometrics*, *9*(1).

Jaffe, J. F. (1974). Special information and insider trading. *The Journal of Business*, *47*(3), 410-428.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, *23*(2), 389-416.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430-454.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kaniel, R., Saar, G., & Titman, S. (2005). Individual Investor Trading. *Working Paper.*

Karolyi, G. A. (2001). Why stock return volatility really matters. *Strategic Investor Relations*.

Kendall, M. G., & Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, *116*(1), 11-34.

Keynes, J. M. (1937). The general theory of employment. *The Quarterly Journal of Economics*, *51*(2), 209-223.

Khalili Araghi, M., & Mohazzab Pak, M. (2013). Assessing the Exchange Rate Fluctuation on Tehran's Stock Market Price: A GARCH

Application. *International Journal of Management and Business Research (IJMBR)*, *2*(2), 95-107.

Kim, D., & Kon, S. J. (1994). Alternative models for the conditional heteroscedasticity of stock returns. *Journal of Business*, 563-598.

Kleidon, A. W. (1986). Variance bounds tests and stock price valuation models. *The Journal of Political Economy*, 953-1001.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, *12*(3), 445-475.

Kothari, S. P., Shanken, J., & Sloan, R. G. (1995). Another look at the cross-section of expected stock returns. *The Journal of Finance*, *50*(1), 185-224.

Kotiaho, H. (2010). Idiosyncratic risk, financial distress and the cross section of stock returns. *Working Paper.*

Kumar, A., & Lee, C. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, *61*(5), 2451-2486.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315-1335.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, *49*(5), 1541-1578.

Lamoureux, C.G. & Lastrapes, W.D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance* 45, 221–229.

Latoeiro, P., Ramos, S. B., & Veiga, H. (2013). Predictability of stock market activity using Google search queries. *Working Paper 13-06*. Universidad Carlos III de Madrid.

Lease, R. C., Lewellen, W. G., & Schlarbaum, G. G. (1974). The individual investor: attributes and attitudes. *The Journal of Finance*, *29*(2), 413-433.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance, 26, 2277-2299.*

Lee, C., & Swaminathan, B. (2000). Price momentum and trading volume. *The Journal of Finance, 55*(5), 2017-2069.

Lee, C., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The Journal of Finance*, *46*(1), 75-109.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, *26*(12), 2277-2299.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, *105*(1), 1-28.

Lemmon, M. & Portniaguina, E. (2006). Consumer confidence and asset prices: some empirical evidence. *Review of Financial Studies, 19, 1499-1529.*

LeRoy, S. F. (1973). Risk aversion and the martingale property of stock prices. *International Economic Review*, *14*(2), 436-446.

LeRoy, S. F., & Porter, R. D. (1981). The present-value relation: Tests based on implied variance bounds. *Econometrica: Journal of the Econometric Society*, 555-574.

Liesenfeld, R. (2001). A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics*, 104, 141-178.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13-37.

Llorente, G., Michaely, R., Saar, G., & Wang, J. (2002). Dynamic volume-return relation of individual stocks. *Review of Financial studies*, *15*(4), 1005-1047.

Lo, A. (2007). Efficient markets hypothesis. To appear in L. Blume and S. Durlauf, The New Palgrave: A Dictionary of Economics, Second Edition, 2007. New York: Palgrave McMillan.

Lobato, I., & Velasco, C. (2000). Long memory in stock market trading volume. *Journal of Business and Economic Statistics,* 18, 410–427.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, *17*(1), 59-82.

Malkiel, B. G., & Xu, Y. (2002). Idiosyncratic risk and security returns. *University of Texas at Dallas (November 2002).*

Malkiel, B.G. (1973). *A Random Walk Down Wall Street*.W.W. Norton & Co., New York.

Malkiel, B. G. (2003). *A random walk down Wall Street: the time-tested strategy for successful investing*. WW Norton & Company.

Mandelbrot, B., (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394–419.

Mangani, R. (2009). Macroeconomic effects on individual JSE Stocks: a GARCH representation. *Investment Analysts Journal*, *69*, 47-57.

Markowitz, H. (1952). Portfolio selection*. *The Journal of Finance*, *7*(1), 77-91.

Markowitz, H. (1959). *Portfolio selection: efficient diversification of investments* (No. 16). Yale University Press.

Marsh, T. A., & Merton, R. C. (1986). Dividend variability and variance bounds tests for the rationality of stock market prices. *The American Economic Review*, *76*(3), 483-498.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, *15*(2), 145-161.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483-510.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 768-783.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, 315-335.

Neal, R., & Wheatley, S. M. (1998). Do measures of investor sentiment predict returns?. *Journal of Financial and Quantitative Analysis*, *33*(4).

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175.

Odean, T. (1998). Do investors trade too much?. *Available at SSRN 94143*.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance*, *53*(6), 1887-1934.

Omran, M. F., & McKenzie, E. (2000). Heteroscedasticity in stock returns data revisited: volume versus GARCH effects. *Applied Financial Economics*, *10*(5), 553-560.

Osborne, M. F. (1959). Brownian motion in the stock market. *Operations research*, *7*(2), 145-173.

Park, J., Konana, P., Gu, B., Kumar, A., & Raghunathan, R. (2010). Confirmation bias, overconfidence, and investment performance: Evidence from stock message boards. *McCombs Research Paper Series No. IROM-07-10*.

Pascal, B. (1670).Pensées. éd. L. Brunschvicg/CM Granges (Paris 1964), 108 f.: 129.

Pouget, S., & Villeneuve, S. (2012). A Mind is a Terrible Thing to Change: Confirmation Bias in Financial Markets (No.720). *Institut d'Économie Industrielle (IDEI),* Toulouse.

Preis, T., Moat, H. S., Stanley, H. E., & Bishop, S. R. (2012). Quantifying the advantage of looking forward. *Scientific reports*, *2*.

Qiu, L. & Welch, I. (2005) Investor Sentiment Measures. *Working Paper,* Brown University.

Rabin, M., & Thaler, R. H. (2001). Anomalies: risk aversion. *The Journal of Economic Perspectives*, *15*(1), 219-232.

Rangaswamy, A., Giles, C. L., & Seres, S. (2009). A strategic perspective on search engines: Thought candies for practitioners and researchers. *Journal of Interactive Marketing*, *23*(1), 49-60.

Rao, A. (2008). Analysis of Volatility Persistence in Middle East Emerging Equity Markets. *Studies in Economics and Finance 25*(2), 93–111.

Redding, L.S. (1996). Noise Traders and Herding Behavior. *IMF Working Paper,* 104.

Reinganum, M. R. (1983). The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects. *Journal of Financial Economics*, *12*(1), 89-104.

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of financial economics*, *4*(2), 129-176.

Roll, R. (1988). The international crash of October 1987. *Financial Analysts Journal,* 19(35).

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory*, *13*(3), 341-360.

Routledge, B. R., & Zin, S. E. (2010). Generalized disappointment aversion and asset prices. *The Journal of Finance, 65*(4), 1303-1332.

Ruan, T., Sun, Q., & Xu, Y. (2010). When Does Idiosyncratic Risk Really Matter?. *Available at SSRN 1561262*.

Samuelson, P. A. (1963). Risk and uncertainty: A fallacy of large numbers. *Scientia*, *98*(4-5), 108-113.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review*, *6*(2), 634-651.

Schmeling, M. (2009). Investor sentiment and stock returns: some international evidence. *Journal of Empirical Finance*, *16*(3), 394-408.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Schwert, G. W. (1989). Why does stock market volatility change over time?.*The Journal of finance*, *44*(5), 1115-1153.

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, *14*(5), 590-610.

Seyhun, H. N. (1998). *Investment Intelligence from Insider Trading*. Cambridge, Mass.: MIT Press

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of market equilibrium under conditions of risk. *The journal of finance*,*19*(3), 425-442.

Shefrin, H. (1999). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press.

Shefrin, H. (2000). *Börsenerfolg mit behavioral finance*. Schäffer-Pöschel.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, *40*(3), 777-790.

Shiller, R. J. (1979). The volatility of long-term interest rates and expectations models of the term structure. *The Journal of Political Economy*, 1190-1219.

Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review 71*,421-436.

Shiller, R. J. (1981). The Use of Volatility Measures in Assessing Market Efficiency. *The Journal of Finance, 36*(2), 291-304.

Shiller, R. J. (1990). Market volatility and investor behavior. *The American Economic Review*, *80*(2), 58-62.

Shiller, R. J. (2000). Measuring bubble expectations and investor confidence.*The Journal of Psychology and Financial Markets*, *1*(1), 49-60.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives*, *17*(1), 83-104.

Shiller, R. J., Fischer, S., & Friedman, B. M. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*, *1984*(2), 457-510.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives*, *4*(2), 19-33.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, *52*(1), 35-55.

Sias, R. W. (2004). Institutional herding. *Review of Financial Studies*, *17*(1), 165-206.

Sias, R. W., Starks, L. T., & Tiniç, S. M. (2001). Is noise trader risk priced?.*Journal of Financial Research*, *24*(3), 311-330.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, *69*(1), 99-118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129.

Swaminathan, B. (1996). Time-varying expected small firm returns and closed-end fund discounts. *Review of Financial Studies*, *9*(3), 845-887.

Solt, M.E., & Statman, M. (1988). How useful is the Sentiment Index?. *Financial Analyst Journal*, *44*, 45-55.

Suhoy, T. (2009). *Query indices and a 2008 downturn: Israeli data*. Technical report, Bank of Israel.

Tauchen, G. E., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485-505.

Taylor, S. (1986). *Modelling Financial Time Series*, Wiley, New York.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139-1168.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39-60.

Thaler, R. (1993). Advances in Behavioral Finance. *New York: Russell Sage Foundation*.

Thaler, R. H. (1999). The end of behavioral finance. *Financial Analysts Journal*, 12-17.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, *76*(2), 105.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, S251-S278.

Tversky, A., Kahneman, D., & Choice, R. (1981). The framing of decisions. *Science*, *211*, 453-458.

Verma, R., Baklaci, H., & Soydemir, G. (2008). The impact of rational and irrational sentiments of individual and institutional investors on DJIA and S&P500 index returns. *Applied Financial Economics*, *18*(16), 1303-1317.

Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, *36*(6), 1808-1821.

Von Neumann, & Morgenstern, O.J. (1947). Theory of games and economic behavior. (2nd Ed.) Princeton: Princeton University Press.

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting*, *30*(6), 565-578.

Wei, S. X., & Zhang, C. (2006). Why Did Individual Stocks Become More Volatile?*. *The Journal of Business*, *79*(1), 259-292.

Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management*, *26*(3), 12-17.

Xu, Y., & Malkiel, B. G. (2003). Investigating the Behavior of Idiosyncratic Volatility*. *The Journal of Business*, *76*(4), 613-645.

Yavan, Z. A., & Aybar, C. B. (1998). İMKB'de Oynaklık. *İMKB Dergisi*, *2*(6), 35-47.

APPENDIX

GOOGLE TRENDS

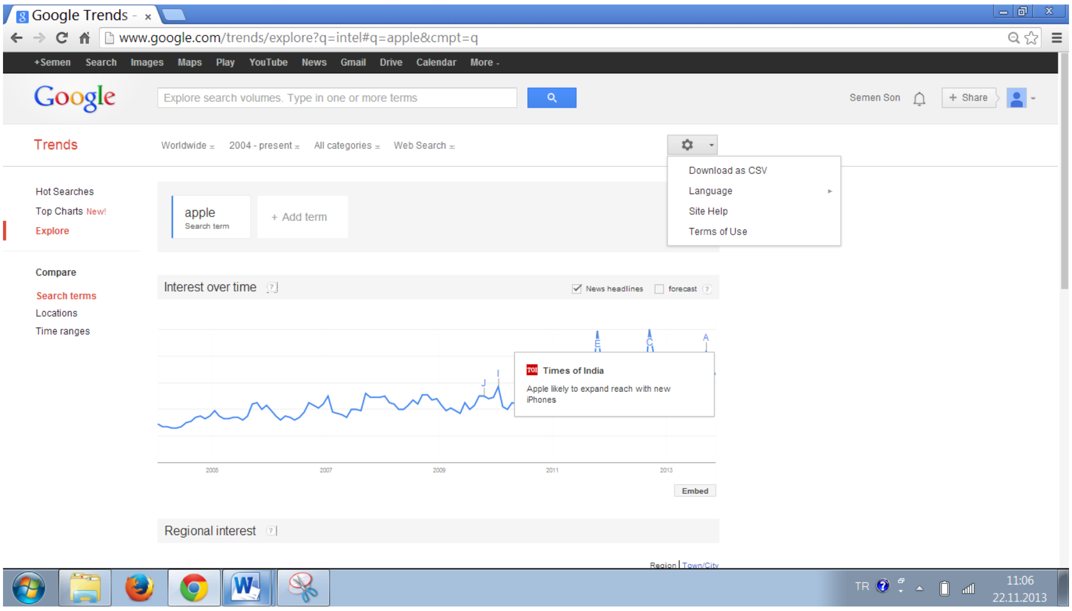

Figure 2 Google Trends Screen Print

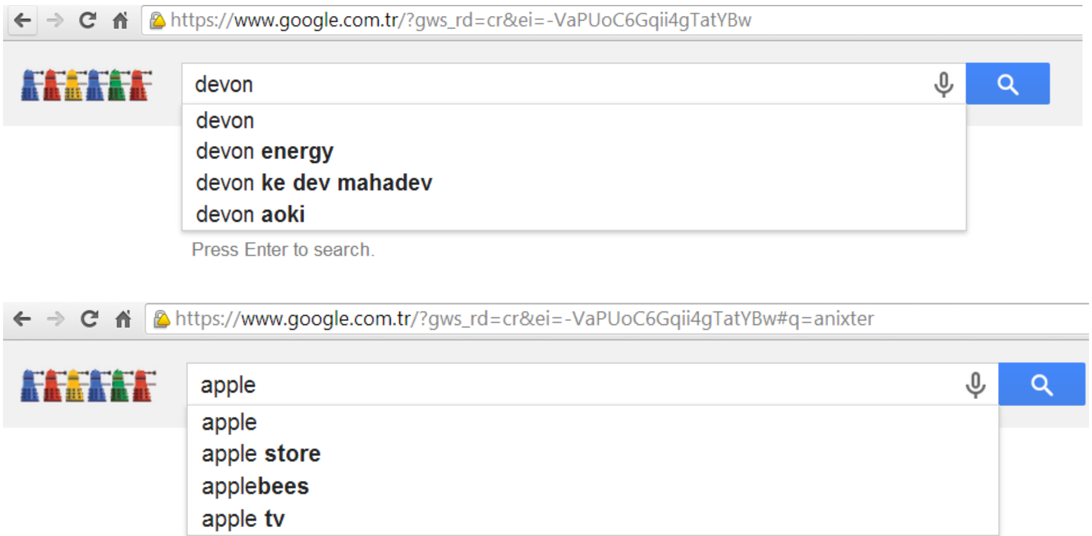

Figure 3 Google Autocomplete Feature

VITAE

Semen Son, holds a BA degree in Political Science and International Relations from Bogazici University and an MBA degree with a concentration in finance and accounting from Pennsylvania State University, The Smeal College of Business Administration. She started her professional career at IS Asset Management in 2002 and worked at companies such as Ford Otosan, REHAU, and Ashmore Asset Management. Her last corporate position was General Manager at Varlik Investment Trust in Istanbul after which she moved back to Izmir in 2010 to work at her family business and commence doctoral studies. Besides her professional career, she has worked as a teaching assistant between 1998-2001 at Penn State. During that time she also completed the curriculum coursework required by the MS in Accounting degree. Other academic experience includes part-time lecturer positions at Penn State, Yeditepe University and Yasar University. Semen is a holder of the Advanced Degree Licence awarded by the Capital Markets Board of Turkey (SPK).