Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

Master's thesis

Trading volume and expected stock returns: a meta-analysis

Author: **Bc. Josef Bajzík**

Supervisor: **doc. PhDr. Tomáš Havránek, Ph.D.**

Academic Year: **2018/2019**

Declaration of Authorship

I hereby proclaim I wrote my master thesis on my own under the leadership of my supervisor and that the included references are all the resources I have used. I also proclaim that I did not use any part of the thesis to earn another university degree.

I grant Charles University permission to reproduce and distribute copies of this thesis document in whole or in part.

Prague, May 8, 2019

Signature

Acknowledgments

I would like to express my gratitude to doc. PhDr. Tomáš Havránek Ph. D. for pushing me to my best and beyond in the area of meta-analysis. His perspective from the top of things inspires me and his professionalism and insight into econometric help me to go behind the theory to experience the practice of the economic research. Moreover, I want to thank Mgr. Anton Astakhov for his valuable points to the topic at the very beginning of the data collection.

Abstract

I investigate the relationship between expected stock returns and trading volume. I collect together 522 estimates from 46 studies and conduct the first meta-analysis in this field. Use of Bayesian model averaging and Frequentist model averaging help me to discover the most influential factors that affect the return-volume relationship, since I control for more than 50 differences among primary articles such as midyear and type of data, length of the primary dataset, size of market, or model employed. In the end, I find out that the relation between expected stock returns and trading volume is rather negligible. On the other hand, the contemporaneous relation between returns and volume is positive. These two findings cut the mixed results from previously written studies. Moreover, the investigated relationship is influenced by the size of country of interest and the level of its development. Besides the primary studies that employ higher data frequency provide substantially larger estimates than the studies with data from longer time periods. On the contrary, there is no difference among different estimation methodologies used. Finally, I employ classical and modern techniques such as stem-based methodology for publication bias detection, and I find evidence for it in this field.

JEL Classification	F14, F29, G10, G12, G14, G23
Keywords	Expected stock returns, trading volume, meta-analysis, Bayesian model averaging, publication bias

Title	Trading volume and expected stock returns: a meta-analysis

Author's e-mail	`bajzik22@seznam.cz`
Supervisor's e-mail	`tomas.havranek@ies-prague.org`

Abstrakt

Zkoumám vztah mezi očekávaným výnosem akcií a jejich obchodovaným objemem. Nasbíral jsem dohromady 522 pozorování z 46 studií a provedl jsem první meta-analýzu v této oblasti. Použití bayesovské metody průměrování modelů a frekventistického přístupu průměrování modelů mi pomohlo objevit nejvíce vlivné faktory, které ovlivňují vztak mezi výnosy a obchodovaným objemem, protože jsem obsáhl více než 50 rozdílů mezi primárními články jako průměrný rok a typ datasetu, délka datasetu, velikost trhu a užitého modelu. Nakonec jsem zjistil, že vztah mezi očekávaným výnosem investic a obchodovaným objemem je spíše zanedbatelný. Na druhou stranu, vztah mezi součaným výnosem akcie a obchodovaným objemem je kladný. Tyto dva závěry třídí smíšené závěry z dříve psaných primárních studií. Navíc zkoumaný vztah je ovlivněný velikostí dané země a stavu jejího rozvoje. Kromě toho primární studie využívající vyšší frekvenci sbírání dat poskytují soustavně vyšší odhady než studie s daty z delších časových period. Na druhou stranu není rozdíl mezi výsledky na základě využití různých metodologií. Nakonec jsem použil klasické a moderní metody jako například trychtýřovou metodu ke zkoumání publikačního vychýlení a našel jsem pro něj dostatek důkazů v této oblasti výzkumu.

Klasifikace JEL	F14, F29, G10, G12, G14, G23
Klíčová slova	Očekávaný výnos akcií, obchodovaný objem, metaanalýza, bayesovská statistika, publikační vychýlení

Název práce	Obchodovaný objem a očekávaný výnos akcií: meta-analýza
E-mail autora	`bajzik22@seznam.cz`
E-mail vedoucího práce	`tomas.havranek@ies-prague.org`

Contents

List of Tables

List of Figures

Acronyms

BMA Bayesian Model Averaging

GMM Generalized Method of Moments

FMA Frequentist Model Averaging

IV Instrumental Variable

MRA Meta-Regression Analysis

OLS Ordinary Least Squares

PIP Posterior Inclusion Probability

PMP Posterior Model Probability

SE Standard Error

UIP Uniform Information Prior

WLS Weighted Least Squares

VAR Vector Autoregressive Model

Master's Thesis Proposal

Author	Bc. Josef Bajzík
Supervisor	doc. PhDr. Tomáš Havránek, Ph.D.
Proposed topic	Trading volume and expected stock returns: a meta-analysis

Research question and motivation For the purposes of improving forecasts of returns and return volatility in dynamic context is the relationship between stock market trading volume and returns of seminal importance. The return-volume relationships are of general interest, since they reveal dependencies that can form the basis of profitable trading strategies. The implications for market efficiencies might be formed from it as well. For instance, technical analyst gives less significance to a price increase with low trading volume than to similar price improvement with substantial volume. During last five decades the topic has attracted many researchers who produced abundance of empirical estimates of this relationship. Despite the fact that the majority found positive relationship, the results are not so persuasive and vary broadly.

One of the systematic methods how to make use of all this work is so called meta-analysis (Stanley, 2012). It consists in collecting and summarizing quantitatively all the estimates. It had been used in economics especially last 30 years. This method was used in economics, for example, by Havranek (2010) on the trade effect of currency unions, Horvathova (2010) on the impact of environmental performance on corporate financial performance or by Card & Krueger (1995) on employment effects of minimum wage.

In this area of economics no meta-analysis has been conducted. There are just several comprehensive articles, which summarize previously written works. Among these comprehensive article might be counted, among others, Karpoff (1987), Mahajan & Singh (2009), or Akpansung & Gidigbi (2015). However, they do not take into account in their comparisons the estimation methods used by other authors, or models employed. Moreover, no one corrected for publication bias. It is well-known that publication bias can seriously affected the estimates. Usually, the estimates inconsistent with the theory are repressed. For instance, Havranek & Irsova (2017) contend with the issue of publication bias in case of border effect.

Hypotheses

Hypothesis #1: The literature estimating trading volume and expected stock returns is affected by publication bias.

Hypothesis #2: The publication bias overstates the mean of realized relationship.

Hypothesis #3: The extent of publication bias decreases in time.

Methodology The first step of meta-analysis is the collection of primary studies. I will examine all articles related to the topic. Since the most recent summarizing article is written in 2015 (Akpansung, Gidigbi, 2015) and ends its overview in 2008, I will additionally search the Google Scholar and Repec Ideas for new studies published. On the obtained dataset I will conduct meta-regression analysis (MRA). MRA is the statistical analysis of previously reported, or published research findings on a given empirical effect, or phenomenon. It is a systematic review of all relevant researches about a specific topic. For using modern meta-analysis method and correction for publication bias, I need the standard error of each elasticity (or another statistical tool from which standard error could be computed). Therefore, I will exclude studies, which not cover them.

After collecting whole dataset of estimates and their related differences, such as standard error, model used, estimation method, midyear of data, type of data, level of aggregation etc., I will run Bayesian model averaging on the data set. Hence I will interpret what and how influence the estimates. After that I will employ funnel graphs, chronological ordering of data, or summary statistics to explain the data in more agreeable way to the reader.

When the publication bias is absent the estimates of relationship are randomly distributed around the true mean level of relationship. Nonetheless, some estimates end in the "file drawer" because they are insignificant or have an "incorrect" sign. For instance, if the statistical significance is required, an author who has only few observations may run a specification search until the estimate becomes large enough to compensate the high standard errors. In this specification the regression coefficient corresponding to the standard error shows the magnitude of publication bias and the intercept reveals the magnitude of the relationship corrected for publication bias (therefore, the specification directly addresses hypotheses 1 and 2). Since such a regression is probably heteroscedastic (the random variable is a sample estimate of the standard deviation of the explained variable), in practice it is usually estimated by weighted least squares corrected with the inverse of standard errors took as weights.

In meta-analysis it has to be considered that estimates coming from one study are likely to be dependent. To cope with this problem I employ the Bayesian Model Averaging, which allows for distinction for unobserved inter-study heterogeneity. Inter-study heterogeneity is probably substantial since the primary studies use data from different countries. To address hypothesis 3 I will add an interaction term between the reported standard error and the year of publication of the study. I expect that the magnitude of publication bias to decline in time, which would be in line with the economics-research-cycle hypothesis (Goldfarb, 1995; Stanley et al., 2008).

Expected Contribution I will conduct a quantitative survey of journal articles studying the relationship between trading volume and expected stock returns. In contrary to previous articles written on this topic, I will conduct comprehensive meta-analysis on this topic and take into account publication bias. Until now, no one investigate publication bias in this field. To study publication bias I will use funnel-asymmetry test and precision effect test. When I correct for the publication bias, I expect to obtain estimates of the relationship lower then usually reported. The results can be directly used by traders on the stock exchanges.

Outline

1. Motivation: There is no meta-analysis on relationship between trading volume and expected stock return. There are just few articles summarizing previous results, e. g. Mahajan, Singh, 2009. But none of them dealing with publication bias, or explaining how the model or estimation technique used affects the results. Since it is already shown that publication bias distorts most areas of empirical economics, there is a chance it will be present here as well.

2. Studies on trading volume and expected stock returns: I will describe how people estimate the relationship between trading volume and expected stock returns.

3. Data: I will clarify how I will collect estimates from studies on trading volume and expected stock returns.

4. Methods: I will explain and use meta-analysis methods, such as funnel graphs, chronological ordering of data, summary statistics, or regressions. Moreover, I will use fixed-effects model for deeper robustness check.

5. Results: I will discuss my robustness checks and regressions.

6. Concluding remarks: I will summarize my findings and present their benefits for future researches.

Core bibliography

Akpansung, A. O. & Gidigbi, M. O. (2015). The Relationship between Trading Volumes and Returns in the Nigerian Stock Market. *International Research Journal of Finance and Economics*, 132.

Brennan, M. J., et al. (1998) Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, 49.3: 345-373.

Brennan, M. J. & Subrahmanyam, A. (1995): Investment analysis and price formation in securities markets. *Journal of Financial Economics*, 1995, 38.3: 361-381.

CHOI, K.-H. & KANG, S. (2013): S. Relationship between Stock Returns and Trading Volume: Domestic and Cross-Country Evidence in Asian Stock Markets. *International Conference on Economics and Business Administration*, Busan, Korea, p. 33-39.

CHORDIA, T., ET AL. (2001): Trading activity and expected stock returns. *Journal of Financial Economics*, 59.1: 3-32.

GALLANT, A. R., ET AL. (1992): Stock prices and volume. *The Review of Financial Studies*, 5.2: 199-242.

GOLDFARB, R.S. (1995): The economist-as-audience needs a methodology of plausible inference. *J. Econ. Methodol.*, 2(2), 201–222.

HAVRANEK, T. & IRSOVA Z. (2017): Do borders really slash trade? A meta-analysis. *IMF Economic Review*, 65.2: 365-396.

HORVATH, R., ET AL. (2017): Financial development, rule of law and wealth inequality: Bayesian model averaging evidence.

LEE, B.-S. & RUI, O. M. (2002): The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence. *Journal of Banking & Finance*, 26.1: 51-78.

KARPOFF, J. M. (1987): The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis*, 22.1: 109-126.

MAHAJAN, S. & SINGH, B. (2009): The empirical investigation of relationship between return, volume and volatility dynamics in Indian stock market. *Eurasian Journal of Business and Economics*, 2.4: 113-137.

EL-QIDREH, K. (2013): The relationship between trading volume and stock prices: an empirical study on the top largest listed public corporations in the NASDAQ Exchange of USA.

STANLEY, T.D. & DOUCOULIAGOS, H. (2012):. Meta-regression analysis in economics and business. *Routledge*, 2012.

STANLEY, T.D., ET AL. (2008): Meta-regression analysis as the socio- economics of economics research. *J. Socio-Econ.*, 37(1), 276–292.

Author Supervisor

Chapter 1

Introduction

I conduct a meta-analysis of the relationship between expected stock returns and trading volume. This relationship is of general importance. It exposes dependencies that might form the basis of profitable trading strategies (Chen *et al.* 2001). Moreover, according to Karpoff (1987) there are several reasons, why to be interested in it. Basically, it provides insight into the financial markets' structure itself and it is seminal for event studies. And naturally, there are significant implications of return-volume relations in future markets researches.

The first studies in area of price-volume relationships were written by Granger & Morgenstern (1963) and Godfrey *et al.* (1964). In the next decades several other articles appeared in US, for example Crouch (1970) and Jain & Joh (1988). Later on researches from every continent started investigating return-volume relations and the studies related to the price-volume relations and similar topics flooded the world. For instance, Lo & Wang (2000) two decades ago found 190 articles studying price-volume relationship from different perspectives. Therefore, I need to be more specific about my focus to be capable to perform meta-analysis. I decide to investigate the expected stock returns and trading volume relationship, since this is not investigated in depth as is other similar topics are.

There were written just several summarizing articles on this topic. I have already shared Karpoff (1987) and among others I can mention, for instance Mahajan & Singh (2009a) and Akpansung & Gidigbi (2015). In short, they provide an overview of currently written literature related to the topic and comment on the major results such as there is mixed evidence for both relationships between expected stock returns and trading volume and between trading volume and stock returns. From previously written summaries one can, for example, say that the level of data aggregation does not have clear evidence of the estimated effect. But no one up till now has tried to figure out why the results differ and whether the publication selection is present between the results or not. Therefore, I decide to investigate the primary studies more in depth via meta-analysis which comprehends all written literature.

Naturally, the first step is the collection of primary studies using RePEc Ideas and Google Scholar. Besides, for each estimated coefficient I need standard error, t-statistic, or p-value. Then I gather data about type of data, midyear of data, number of observation, data frequency, or estimation methodology among other figures. Overall I collected 522 observations from 46 studies.

First step in my analysis is investigation of publication bias. I use funnel plots, OLS, between effects, WLS, IV and some modern approaches like stem-based approach. These are introduced recently since the common formal tests are often questioned by the public. Similar approaches are used, for example, by Havranek & Sokolova (2019). After thorough investigation of publication bias I come to two convincing results regardless of technique employed. I discover that the size effect has negligible value and is insignificant when it is corrected for the publication bias. The second finding from publication bias investigation shows that there is evidence for presence of the publication selection bias in the data. I also find that the phenomena of publication bias is not decreasing in time and that the publication selection is not affected by the journal quality.

The comparison of primary articles studying expected stock returns and trading volume is not so straightforward. For instance, the econometric approach evolves with proceeding of time. At the beginning, Crouch (1970) uses the price changes and volume. During next decade, the returns have substituted the price change, since the return is more natural measure comparable even among stocks with different a starting price. Moreover, the measure of trade volume has changed. Some of the first authors (e. g. Epps & Epps 1976) used number of shares traded. Later on, number of shares was replaced by dollar share volume (Brennan *et al.* 1998) and nowadays the most studies use turnover as volume measure (Zhong *et al.* 2018). Besides the econometric models developed among others. From simple models through VAR models to Fama-MacBeth approaches employed, for instance, by Chordia *et al.* (2001), which is current workhorse in study of the exchanges.

Therefore, I employ Bayesian model averaging (BMA) and Frequentist model averaging (FMA) to determine the most influential factors in estimating the relationship between trading volume and expected stock returns. The Frequentist model averaging is new technique in the meta-analytical field. It was first used in meta-analysis probably by Havranek *et al.* (2017). These two approaches used weighted averages of the best model, but each employs a different approach. While the BMA is based on Metropolis-Hastings algorithm, the FMA stands on the goodness of fit and parsimony of potentially included models. To enable the model averaging FMA employs orthogonalization of the variable space. Both methods tackle the problem of heterogeneity in the data. First of all, results from BMA and FMA confirmed both major findings from publication bias investigation. Moreover, other results from BMA and FMA are equally interesting. I can confirm that there is small or even negligible relationship between trading volume and expected stock returns. It is in line with the findings from primary studies such as Lee & Rui (2002), or Gurgul *et al.* (2007) and from summarizing articles Mahajan & Singh (2009a), Akpansung & Gidigbi (2015), but these results contradicts findings from Hafner (2005), or Hu (1997). Similarly, the conclusion for the contemporaneous relationship between stock returns and trading volume is in line with some primary articles (Lee & Rui 2000), but stands against others (Chordia *et al.* 2001). The estimated coefficient in contemporaneous case is about 0.128 units higher than in case of expected stock returns. This result is statistically significant at 1% level. Besides these major findings, other interesting facts arise. For example, in bank sector, which is investigate by Al-Jafari & Tliti (2013) and Rotila *et al.* (2015), I find negative and significant effect on the relationship trading volume and expected stock returns. On the contrary the higher the frequency of the data collection in the primary studies the higher the estimated coefficient. Moreover, there is no difference in estimations caused by different level of aggregation used in primary studies. The same is true for type of data, because all three types; time series, cross-sectional and panel data produce comparable results. On the other hand, the results for developing and smaller countries are substantially lower (-0.077 and -0.051, both significant at 1% level) than the results for developed and large markets. The last interesting finding comes from the research among studies published in impacted journals. The results published in more influential journals are lower than other estimates on average. The direction of this finding correspondents with the one from publication bias section, but in that case it was insignificant. Thus, one should be vigilant to do any impetuous conclusions based on this evidence. All these results might be used as a baseline in model calibration or directly in traders' strategies.

The structure of the thesis is as follows: Chapter 2 summarizes all previous findings in area of return-volume relationship. In Chapter 3 the data collection approach is explained and variables used are enumerated. The following Chapter 4 discusses my approaches to publication bias investigation, e. g. with funnel plots, OLS, WLS, IV, BE, or with the most modern techniques. Besides, the BMA methodology is described, and FMA approach is explained in a nutshell. Finally, Chapter 5 explains the results of publication bias and discusses the conclusions from BMA and FMA models. The summary statistics are used to sketch the results in more agreeable way to the reader. In the last chapter, Chapter 6, the findings are summarized.

Chapter 2

Literature Review

The researchers are attracted by trading volume for years and it results in voluminous supply of literature on this topic. For example, even Lo & Wang (2000) two decades ago found almost two hundreds of articles relating to trading volume from various fields - finance, economics or accounting among others. And during last two decades many more new articles have been published. My main focus is on those articles in the financial field. Besides volume return relations financial researches discuss mainly volume-return volatility relations, market microstructure, and models of asymmetric information Brandle (2010). As usually, the majority of studies in this area originates from US stock markets (e. g. Granger & Morgenstern 1963 or Epps & Epps 1976). In addition to this, numerous literature over the recent years approaches from emerging markets (e. g. De Meiros & Van Doornik 2008 or Tapa & Hussein 2016).

2.1 Importance of Volume-Return Relationship

The expected return-volume relationships are of general interest, since they reveal dependencies that can form the basis of profitable trading strategies (Chen *et al.* 2001). Return might be interpreted as the evaluation of new information. On the other hand, the volume can indicate the level of disagreement about the evaluation of the new information among the investors (Mahajan & Singh 2009a). According to Karpoff (1987) there are at least four reasons of importance of price-volume relation. First, it provides insight to the structure of financial markets itself. This means that the return-volume relations depend on a rate of information flow to the market, how the information is spread, it sheds light to the existence of short sales constraints or to the extent to which market prices carry the new information. Second, the price-volume relation is important to the discussion over the empirical distribution of speculative prices. Third, the price-volume relation is crucial for event studies. These studies draw inferences from combination of price and volume data. The power of such tests increases when incorporated price changes and volume are jointly determined. For instance, Richardson *et al.* (1986) investigate price changes and trading volume to test of occurrence of dividend clienteles. And fourth, there are significant implications of price-volume relations for researches into futures markets. The volume of traded futures contracts is affected by the time to deliver (Grammatikos & Saunders 1986) and by the price variability (Cornell 1981).

Hence, the relation between return and volume has obtained substantial attention to better the understanding of the microstructure of the stock markets. Besides this relation clarifies the stock markets efficiency (Mahajan & Singh 2009a). Based on the objective of this thesis, the focus of following literature review is placed on expected return-volume relations and other studies (e. g. those on volume-return volatility relations) are out of scope of this research.

2.2 Contemporaneous Volume-Return Relation

Much of the early research discussed the contemporaneous relationship between volume and returns. The seminal work in this area is Karpoff (1987). His review article inspects two "stylized facts". First of them is that there is existence of a positive relationship between volume and the price change *per se*. And the second one is that positive relationship between volume and the absolute value of the price change exists. Those two "stylized facts" are confirmed by newer studies usually, but more than few authors contradict it (e. g. Sheu *et al.* 1998) and this is one of the reasons why I decided to conduct a meta-analysis in this field.

2.2.1 Correlation between Volume and Returns

A detailed analysis of return-volume dynamics is crucial to gain knowledge of issues relating to information flow in the market and market efficiency (Mahajan & Singh 2009a). The contemporaneous relationship between returns and volume clarifies information regarding asymmetry of trading volume in the markets. As mentioned from Karpoff (1987) the positive relationship between return and volume is widely acknowledged in the financial literature. The positive relationship is observed in stock and bond markets, but it is not observed in futures markets (Chen *et al.* 2001).

The positive contemporaneous price-volume relation is supported for instance by Epps & Epps (1976), who studies twenty individual US stocks in January 1971.

He uses number of transactions as a measure of volume. The same is confirmed by Harris (1987), who uses besides the number of transactions for volume the daily data for individual stocks. Wood *et al.* (1985), on the other side find the relationship negative and Gurgul *et al.* (2007) who study individual stocks on German DAX index find no relationship at all. The positive results is supported by articles using longer measures of trading volume and returns. For instance, it is supported by Lee & Swaminathan (2000), who used monthly data on individual stocks and share turnover as a measure of volume. The same type of data (but gathered annually) and similar results exhibits Comiskey *et al.* (1987). The positive contemporaneous relation is moreover confirmed by results from aggregate stock markets. On S&P 500 index data Jain & Joh (1988) confirm the positive relation. The same indicates Lee & Rui (2000) on all four China's stock markets and Lee & Rui (2002) on all three stock markets in New York, Tokyo and London.

These findings about stock markets might be supported by the model suggested by the Jennings *et al.* (1981), who extended for short selling the sequential information arrival model introduced by Copeland (1976). In general, short sales are generally more costly than long positions. It restricts ability of some investors to trade on new information. Based on these indices, Jennings *et al.* (1981) show that the volume is higher when previously uninformed trader interprets new information optimistically than when the trader is a pessimist. Since the prices decreases with the pessimist who sells and increases with optimists who buys, it is then implied that volume is higher when price increases and low when the price decreases. The sequential information arrival model, moreover, suggest a dynamic relationship, where lagged values of trading volume may have the ability to predict current absolute returns, and vice-versa (Darrat *et al.* 2003). This hypothesis is moreover supported by empirical findings from futures markets, where costs of taking short and long positions are symmetric. Among such articles belong, for example, Kocagil & Shachmurove (1998) or Mcmillan & Speight (2002).

From the newest studies, one cannot forget the article investigating the Vietnamese stock market written by Vo (2017). He focuses on trades of foreign investors on Ho Chi Minh City stock exchange and finds out that purchases of foreign investors are negatively contemporaneously correlated with stock returns. On the other hand, the sales of foreign investors are positively correlated.

2.2.2 Correlation between Volume and Absolute Returns

As indicates an old saying on Wall Street "It takes volume to make prices move" (Karpoff 1987), there is positive correlation between volume and absolute returns.

It has been strongly confirmed by prior research as well. Among the studies, there arise two alternative explanations for this phenomenon, mixture of distributions hypothesis and the sequential information arrival model (Brandle 2010), which was already mentioned in Subsection 2.2.1.

Unlike to sequential information arrival model; the mixture of distribution hypothesis suggests only a contemporaneous relationship between volume and absolute returns (Mahajan & Singh 2009a). In a nutshell, the observation that the price changes of speculative assets appear to be symmetrically distributed and uncorrelated with each other is at the outset. On the other hand, the distribution is in fact kurtotic relative to the normal distribution. For example, Epps & Epps (1976) develop one mixture of distributions hypothesis model. In this model, the variance of the change in price on single transactions is conditional on the volume of the same transaction. Consequently, these models imply contemporaneous positive correlation between absolute returns and volume. Besides Epps & Epps (1976), these models are associated, for instance, with Clark (1973), Tauchen & Pitts (1983), or Harris (1987).

The prior research unanimously confirms the positive correlation between absolute returns and volume. This conclusion is drawn regardless of data frequency, aggregation level and time period of the sample. Crouch (1970) finds a positive correlation on daily price changes and volumes for both market and individual levels. The findings of Comiskey *et al.* (1987) confirmed the hypothesis of mutual positive correlation between volume and absolute returns even in yearly data on individual stock. Jain & Joh (1988) find the hypothesized correlation on completely opposite side of data spectrum - using one-hour intervals and market index data. In the middle of Comiskey *et al.* (1987) and Jain & Joh (1988) stands Lee & Swaminathan (2000) studying monthly data over 30 years time period, again with the same conclusions.

The conclusions from the latest studies confirm the same. Ciner (2002) finds a positive and significant relationship between trade volume and absolute returns on the Toronto Stock Exchange before and after automation. Ciner (2003) specializes on small firm stocks in US and France and acknowledge the positive relationship at 5% and 10% significance level, respectively. Similarly, Assogbavi *et al.* (2007) when uses data from Russian Stock Exchange from 1997 to 2005 reassure the same conclusion. They report 3 out of 28 results negative, but these are insignificant. And finally, the last study engaged in the relationship between volume and absolute returns published Yonis (2014). His five results for US, Taiwan, Hong Kong, Korea and Singapore are all positive and significant.

2.3 Dynamic Volume-Return Relation

This area investigates the joint dynamics between expected stock returns and trading volume. This investigation is in most cases done by vector autoregressive models (VARs). In brief, these models study pure time-series (on the contrary to the previously discussed approaches) and these are dynamics models. It means that in the same moment the models investigate the influence of expected returns on volume and volume on expected returns as well (Brandle 2010). Since time series are used, several lags of each variable might be used in both equation. For more detailed description of different VARs used among the studies, please refer to section Subsection 3.2.5.

There is very little evidence of time-series relations between lagged trading volume and stock returns, even over different time horizons. For instance, Statman *et al.* (2006) studied roughly 2,000 US stocks, but finds no significant relation between past share turnover and returns. They used monthly data and 40 year sample. Using comparable dataset, Chuang & Lee (2006) agrees with this finding. Similar conclusion draw from daily data of 29 DAX companies Gurgul *et al.* (2007). They find a significant relationship between lagged trading volume and current stock returns only on one case at 5% level. They conclude that this evidence is in line with efficient market hypothesis. From recent studies the same is found by Al-Jafari & Tliti (2013). They study banking sector on Jordanian's Amman Stock Exchange during years 2006 and 2012. One exception is Pisedtasalasai & Gunasekarage (2007) who investigate Asian countries on the turn of the century. Their results are negative in 4 of 5 cases, but insignificant.

Similar findings as for researches on individual stocks comes from studies on aggregate stock markets. Besides Statman *et al.* (2006), this was concluded by Lee & Rui (2000) and Lee & Rui (2002) as well. Some exceptions may be found even in studies on aggregate data. One of them is Devanadhen *et al.* (2010) in their research of Asia-Pacific countries. Their results are rather inconclusive than persuasive. The same can be said about results from Vo (2017), who focuses on Vietnamese stocks.

2.4 Lagged Volume-Return Relationship

Finally, more recent area of volume-return relationship investigation is shortly discussed. It is the most important area of this research. It analyzes whether and how measures of lagged trading volume affect the consecutive returns. Basically, the main question of this research concerns how the different volume levels relate to expected returns.

There are many studies in this research area, for both US and international stock market data. Several authors employ Fama-MacBeth type regressions and their variations. Among such belongs Haugen & Baker (1996). They run regression of stock returns on over 40 firm characteristics such as risk, liquidity, growth potential, price history and price level in the Russel 3000 stock index. They find a negative and statistically significant relationship between the dollar volume-market capitalization ratio and stock returns. Negative relationship between volume level and expected stock returns in the US stock markets confirm Brennan *et al.* (1998) and Chordia *et al.* (2001) as well. In addition to, existence of the negative relationship between volume levels and expected stock returns validates Hu (1997) in the Tokyo stock exchange and by Loukil *et al.* (2010) on Tunisian stock exchange.

In the last five years, many studies that use Fama-MacBeth methodology arose. These are mainly from US and Asia, especially from China. Lewellen (2015) finds the seminal relationship negative on New York Stock Exchange. On the other hand, Chang & Wang (2019) and Han *et al.* (2018) using data for NYSE, AMEX and NASDAQ finds mixed results. On all three US exchanges Chang & Wang (2019) found the relationship mostly slightly negative but insignificant. Han *et al.* (2018) find the relationship negative among overpriced stocks and positive among underpriced stocks. The results for Asia are more unambiguous. All three, Zhong *et al.* (2018) in Asia-Pacific, Long *et al.* (2018) on Chinese Stock Exchange and Yin & Liu (2018) in China, find the relationship always negative. Besides Long *et al.* (2018) find the return-volume relationship significant in all six his cases.

2.5 Other Approaches

There are many approaches to investigation of volume-return relationship, but it is unfeasible to get them all into one single meta-analysis. Therefore, I provide a brief survey of the main approaches and their main conclusions, which cannot be involved in our main research.

2.5.1 Lagged Return-Volume Relationship

Similar VAR approach as to study relation between lagged volume and stock returns is usually used for the inverse relationship. Unlike to results of lagged volume and stock returns, the opposite relationship provides significant results. With these results come, for instance, above mentioned Statman *et al.* (2006), Chuang & Lee (2006), or Lee & Rui (2000). Another example might be Gurgul *et al.* (2007). They find only 1 of 29 relationships between past volumes and stock returns significant and positive at 5% level. On the contrary, they find 22 of 29 relations significant and positive between lagged returns and trading volume.

2.5.2 Granger Causality

I do not include one important feature in the context of VARs, namely Granger-causalities. This kind of causality was proposed by Granger (1969) as "a correlation between the current value of one variable and past values of other variables" (Brandle 2010). Many of the studies, which use VAR models discuss besides the dynamic relationship between volume and returns, Granger-causality as well. One illustration is already mentioned Lee & Rui (2000), others may be found in Mestel *et al.* (2003), or Akpansung & Gidigbi (2015). I do not include these results, since they discuss whether or not volume Granger-cause returns and vice-versa, but I am interested in estimated coefficient between these to variables, not in direction.

2.5.3 Change in Volume

Not all studies used in research of volume-return relationship volume in logs or in levels. Several of them use some measure of change of traded volume in their analysis. First study used such an approach was Ying (1966). He finds that increase in daily trading volume on the NYSE turns into a rise in the price of the S&P 500 Composite Index. Other similar studies followed. Just to mention few of them, e. g. Gervais *et al.* (2001) extend the Ying's analysis or Watkins (2007) finds stocks with high-mean volume growth in the preceding 12 months undergo higher returns over the consecutive one to 60 months. He uses quintile portfolios in his analysis.

2.5.4 Return Autocorrelation

Another area of research, which is out of scope of this study, focuses on the effect of past returns on the current stock returns. For instance, De Bondt & Thaler (1985) find that prior losers outperform prior winners by 25%. It holds for 36 months after portfolio formation. Other result comes from study written by Lee & Swaminathan (2000). They find that the momentum effect is stronger among stocks with higher volume in US exchange markets.

2.6 Review of Articles Included in the Meta-Analysis

The following table shows an overview of all articles that are used for this meta-analysis. It includes the number of observations (n), its means, standard deviations and medians as well.

Table 2.1: Review of Used Articles

Authors	*Country*	*Type*	*n*	*Mean*	*SD*	*Med*
Lee & Rui (2002)	US, Japan, UK	Log-level	9	0.05	0.12	0.03
Mahajan & Singh (2009a)	India	Log-level	2	0.05	0.01	0.05
Ciner (2002)	Canada	Log-log	8	0.01	0.04	0.00
Lee & Rui (2000)	China	Log-level	8	0.24	0.13	0.19
Al-Jafari & Tliti (2013)	Jordania	Log-log	2	0.02	0.02	0.02
Chen *et al.* (2001)	US, Japan, UK, France, Canada, Italy, Swiss, Netherlands, Hong Kong	Level-level	9	-0.01	0.02	-0.01
Ciner (2003)	US, France	Log-log	4	0.02	0.04	0.03
Tripathy (2011)	India	Log-level	2	0.00	0.01	0.00
Kim (2005)	Australia, Japan, Hong Kong, Singapore, US	Loglev	48	0.00	0.04	0.01
Louhichi (2012)	France	Log-level	2	-0.02	0.02	-0.02
Mahajan & Singh (2008)	India	Log-level	1	0.04	N/A	0.04
Tapa & Hussein (2016)	Malaysia	Log-log	2	0.00	0.28	0.00
Tahir *et al.* (2016)	Pakistan	Log-level	6	0.17	0.17	0.16
Saatcioglu & Starks (1998)	Argentina, Brazil, Chile, Colombia, Mexico, Venezuela	Log-level	6	0.09	0.03	0.1
Assogbavi *et al.* (2007)	Russia	Log-level	28	0.13	0.09	0.12
Yonis (2014)	US, Hong Kong, Korea, Singapore, Taiwan	Log-level	20	0.00	0.09	0.01

Continuation of Table 2.1						
Authors	*Country*	*Type*	*n*	*Mean*	*SD*	*Med*
Pisedtasalasai & Gunasekarage (2007)	Jakarta, Malaysia, Philippines, Singapore, Thailand	Log-level	5	-0.01	0.01	-0.01
Rotila *et al.* (2015)	EU	Log-log	12	0.00	0.01	0.00
Le & Mehmed (2009)	Sweden, Denmark, Norway, Finland	Log-level	4	0.01	0.01	0.00
De Meiros & Van Doornik (2008)	Brazil	Log-level	2	0.06	0.02	0.06
Crouch (1970)	US	Level-level	9	0.41	0.19	0.37
McGowan & Muhammad (2012)	Russia	Level-log	2	0.00	0.03	0.00
Sana Hsieh (2014)	Hong Kong, Japan, Malaysia, Philippines, Singapore, Taiwan, Thailand	Log-level	7	0.28	0.09	0.24
Chordia *et al.* (2001)	US	Log-log	36	-0.24	0.06	-0.22
Narayan & Zheng (2010)	China	Log-log	4	0.01	0.03	0.00
Lin & Liu (2017)	US	Log-level	7	0.07	0.06	0.02
Loukil *et al.* (2010)	Tunis	Log-level	1	-0.2	N/A	-0.2
Brennan *et al.* (1998)	US	Log-log	30	0.00	0.00	0.00
Lewellen (2015)	US	Log-level	3	-0.03	0.03	-0.02
Shu *et al.* (2004)	Taiwan	Log-log	4	0.52	0.22	0.56
Hafner (2005)	US	Log-log	6	-0.01	0.06	-0.02
Vo (2017)	Vietnam	Log-log	6	0.02	0.06	0.02
Zhong *et al.* (2018)	Asia-Pacific	Log-level	3	0.00	0.00	0.00
Brandle (2010)	Swiss	Log-log	4	0.03	0.02	0.03
Sheu *et al.* (1998)	Taiwan	Log-log	28	-0.04	0.03	-0.04

Continuation of Table 2.1

Authors	*Country*	*Type*	*n*	*Mean*	*SD*	*Med*
Ochere *et al.* (2018)	NSE	Log-level	1	0.32	N/A	0.32
Long *et al.* (2018)	China	Log-level	5	-0.47	0.03	-0.49
Han *et al.* (2018)	US	Log-level	5	0.00	0.00	0.00
Epps & Epps (1976)	US	Log-log	40	0.2	0.24	0.23
Marshall & Young (2003)	Australia	Log-level	6	-0.15	0.2	-0.13
Chang & Wang (2019)	US	Log-log	28	0.00	0.00	0.00
Yin & Liu (2018)	China	Log-log	4	-0.01	0.00	-0.01
Datar *et al.* (1998)	US	Log-level	35	-0.02	0.01	-0.02
Hu (1997)	Japan	Log-level	62	-0.01	0.01	-0.01
Devanadhen *et al.* (2010)	Australia, India, Japan, New Zealand, Taiwan	Log-log	5	-0.01	0.06	0.00
Mahajan & Singh (2009b)	India	Log-level	1	0.04	N/A	0.04

Notes: This table shows all articles used in the meta-analysis. At each article is mentioned, whether they used log-log, log-level, level-log, or level-level type of data, number of estimation used from the study, their mean, standard deviation and median.

Chapter 3

Data

Since meta-analysis is research based on primary studies, I started to collect the primary studies in July 2018. The search was conducted primarily through Google Scholar[1] and Scopus[2] by following key: trade | trading and volume and "expected stock return" | "stock return" | "price changes". Since I try to have the sample as comprehensive as possible, I updated it in February 2019 and one hard-copy article was delivered in mid-March 2019. I decided to wait as long as possible, since I aim at detecting every distinguishing sign, that can help me to understand the root of differences in results in primary studies. During the data collection I proceed according to Stanley & Doucouliagos (2012) and I collect essential, typical and value-added variables.

3.1 Essential Variables

Among the essential variables I primarily include effect size of the relationship between returns and volume, standard error and sample size. I discuss them more thoroughly in the following paragraphs, since some of them are not so straightforward.

[1]Google Scholar [online]. Infogram: ©2004 [cit. 10. 7. 2018]. Available from: https://scholar.google.cz/.

[2]Scopus [online]. Infogram: ©2019 [cit. 10. 7. 2018]. Available from: https://scopus.com/.

3.1.1 Effect size

In the case of effect size, I have to contend with the fact that several specifications of the left-hand side are used in the literature. First and second are price change and absolute price change, defined as

$$\Delta P = P_t - P_{t-1}, \tag{3.1}$$

and

$$|\Delta P| = |P_t - P_{t-1}|, \tag{3.2}$$

which are used in the oldest articles. In the newer ones, the returns and absolute returns are employed. These are more common in the literature, since they are more comparable among different stocks, firms, or studies. Returns are then defined as:

$$\Delta Ret = ln(P_t) - ln(P_{t-1}) = ln(P_t/P_{t-1}). \tag{3.3}$$

Absolute returns are captured by Equation 3.4 and they are distinguished by additional dummy variable.

$$\Delta ARet = |ln(P_t) - ln(P_{t-1})| = |ln(P_t/P_{t-1})|. \tag{3.4}$$

Moreover some articles use value of volume in logarithms and again other use just raw data. Thus, the observations are not directly comparable. I need to deal it in sake of interpretation of results. In this matter I follow Valickova *et al.* (2015). They use partial correlation coefficients (r). These are commonly employed in economic meta-analysis (e. g. Doucouliagos 2005). The partial correlation coefficients are derived from t-statistic of the estimate and residual degrees of freedom Greene (2008).

$$r_{ij} = \frac{t_{ij}}{\sqrt{t_{ij}^2 + df_{ij}}}, \tag{3.5}$$

where r_{ij} stands for the partial correlation coefficient from the ith estimate of the jth study. The t is the corresponding t-statistic and df are degrees od freedom. Since I am not able to collect degrees of freedom for every estimate, I substitute it with the number of observations related to each particular estimate. The sign of the partial correlation coefficient stays the same as the sign of the original coefficient.

Sometimes the exact number of observations is not directly mentioned. If it is the case when using time series data, I calculate the number of observations as the duration of the data set times the days of the week, since there is no exchange during weekends. It is case, for example, of Mahajan & Singh (2009a), Tripathy (2011), or

Vo (2017). Similarly, I solve the case of Louhichi (2012), who collects time series of 101-five-minutes intervals in a day. I multiply the number of week days of data duration in this article by 101. If the panel data are used in primary studies and the total number of observations is not mentioned, I proceed as in case of time series data, but moreover I multiply number of time series units with number of cross-sectional units. I employ this method, for instance, in case of Brennan *et al.* (1998), or Zhong *et al.* (2018).

The mean and median of the partial correlation coefficient are 0.02 and 0.00, respectively, which ranks the effect size of return-volume relationship among insignificant effects according to Doucouliagos (2011). He developed a guideline for partial correlation interpretation in meta-analysis based on 22 thousand observations from 41 meta-analysis to counterweight Cohen (1988)'s conventional guidelines for zero order correlations. The division of partial coefficient effect sizes according to Doucouliagos (2011) is following: the effect is considered "small" when the partial correlation coefficient ranges between 0.07 and 0.17, "medium" when it lies between 0.17 and 0.33 and "large" when the absolute value is greater than 0.33. The histogram of the partial correlation coefficients is captured in Figure 3.1.

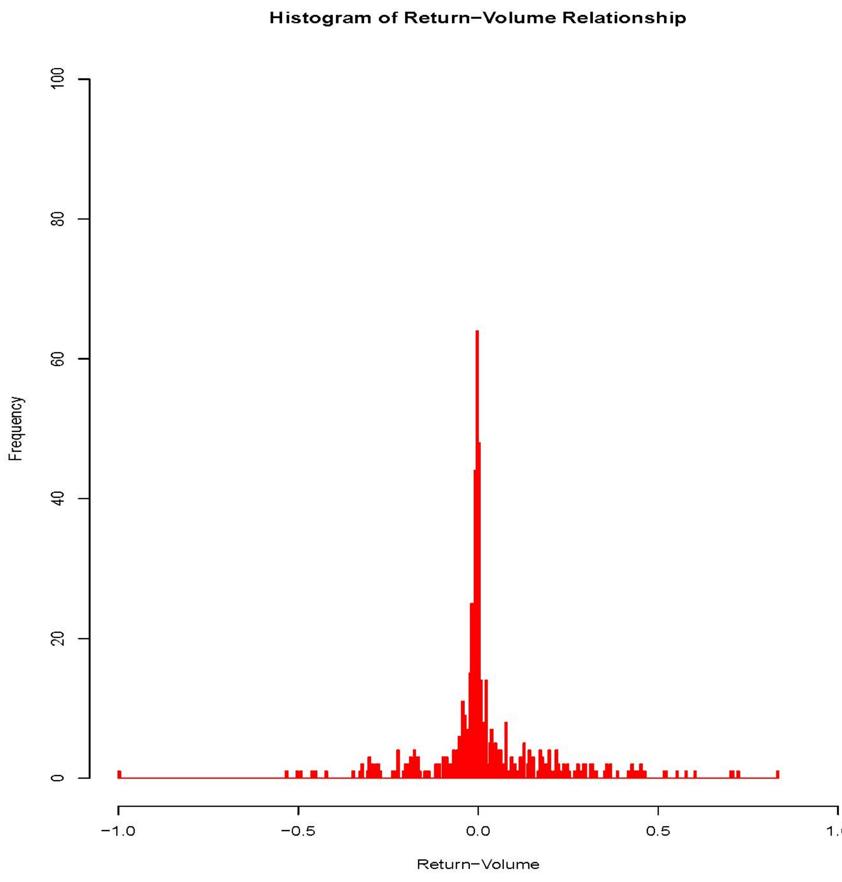

Figure 3.1: Histogram of Return-Volume Relationship

3.1.2 Other Essential Variables

Besides the estimates of the relationship itself, I add two dummy variables whether the case that the relationship is *Contemporaneous* or *Dynamic*. That means, whether the returns and volume are from the same time period, or volume collected is one period behind returns.

Other essential variables are standard error and sample size. When the standard error, or t-statistic, or exact p-value is not mentioned, I do not include such estimates into this meta-analysis. From such estimates I am not able to draw exact inference and discuss the publication bias. If I contend with t-statistic, or p-values instead of standard error, it is recalculated[3] to standard error because I need to compare the inferences.

Since I recalculated the effect size to partial correlation coefficients, I need to adjust standard errors. Again I follow modification adapted by Valickova *et al.* (2015). They adapted following formula from Fisher (1954):

$$SEr_{ij} = \frac{r_{ij}}{t_{ij}}, \tag{3.6}$$

where SEr_{ij} denotes the standard error of the particular partial correlation coefficient r_{ij}. The t_{ij} is the t-statistic from the ith regression of the jth study. The distribution of standard errors is to be found in Appendix A in Figure A.1.

3.2 Typical Variables

Other variables used frequently across the studies are classified by Stanley & Doucouliagos (2012) as "Typical". Among these I can rank return measures, volume measures, research area, information about data set, model used, or estimation methodology. Purpose of these variables is to correct the original econometric research. I include each variable, which could be distinguished in more than 2% of cases, which means it has 11 or more observation. In the following paragraph, I discuss each category a little.

3.2.1 Return measures

As the return-volume relationship is investigated for more than five decades, four measures of returns arose during the time. Since I want to distinguish among these

[3] Stat Trek: T Distribution Calculator: Online Statistical Table [online]. Infogram: ©2016 [cit. 10. 10. 2016]. Available from: http://stattrek.com/online-calculator/t-distribution.aspx.

differences, I capture the different deviations from simple return-volume relationship by dummies. First dummy, *Absolute* is for the case, when the returns in primary study are only in absolute values. Second and third dummies are quite similar. *Abnormal* counts only with return above average return from previous time framework. *Excess* considers only return above risk free rate.

3.2.2 Volume measures

There are several measures that can be used to capture trading volume. At the beginning authors as Crouch (1970), or Epps & Epps (1976) used number of shares traded as volume measure. At the turn of the century, this measure was replaced by dollar share volume (e. g. Chordia *et al.* 2001, or Brennan *et al.* 1998). That means that instead of pure number of shares traded, the value (especially in US dollars) of trade was considered as important. Then Lo & Wang (2000) examined share volume, dollar volume, turnover, number of trades, trading days per year, and contracts traded, and recommend turnover as the most natural measure of trading volume in the stock market. The reason is that it yields the sharpest asset pricing implications. Therefore nowadays, most studies use turnover as volume measure (e. g. Long *et al.* 2018, Chang & Wang 2019, Zhong *et al.* 2018). The turnover is defined as the number of shares traded during a time period divided by the number of shares outstanding at the end of the time period. I capture the differences between the volume measures by dummies and choose the turnover as a benchmark. Besides these three dummy categories, I add one variable to capture, whether the volume series in primary study is *Detrend*ed by linear and quadratic time trend or not.

3.2.3 Research Area

In addition to different volume and return measures I investigate and capture by dummies research area of particular articles. I differentiate seven categories: First of them is *All*, when authors consider all companies traded on particular stock exchange. Examples of such researches are Marshall & Young (2003), or Han *et al.* (2018). The second largest group represented by Chordia *et al.* (2001) and McGowan & Muhammad (2012) employed just some *Index* (e. g. S&P500). Rotila *et al.* (2015) and Al-Jafari & Tliti (2013) focus on more concrete market. They work only with *Bank*'s stocks. On the contrary, Tahir *et al.* (2016) and Datar *et al.* (1998) investigates only *Nonfinancial* sector. Other research areas, as S&P600 by Ciner (2003) do not meet the condition of at least 2% of estimates. All these observations are recorded in other individual sectors (*Otherinv*).

3.2.4 Information about primary data sets

Besides number of observation, I collect following pieces of information for each observation. I care about *Length* of the time period in years and *Midyear* of the data used. I thought about using the year of publication (*Pubyear*) to capture differences in data of publishing, but it was correlated with *Midyear* above 85%, thus I neglected this idea. This correlation is captured in Appendix A in Figure A.6 and Table A.1. Then, I wanted add squares of *Midyear*, but the squares are strongly correlated with the linear terms (correlation above 97%). Thus, I abandoned from this option. The final correlation between linear and squared terms is in Appendix A in Figure A.7 and Table A.2. Based on findings of Schürenberg-Frosch (2015), linear terms should be sufficient.

Next I distinguish, if the country belongs among *Developed* countries or not. As developed country is classified each country from Central and Western Europe, North America, Australia, New Zealand and Japan. Moreover, I divide the studies according to type of data, whether the data are *Cross-sectional*, *Time-series* or *Panel*. In addition to, I sort data frequency as *Hourly* and less, *Daily*, *Weekly* and *Monthly* or more. Finally, I recognize by variable *Spillover*, whether the relationship between returns and volume is connected to one market, or if the author investigates trade volume in one market and returns in other one.

3.2.5 Models

The models employed by researchers are split into ten groups according to their specification. In the base equation of the estimated dataset is used Fama - Macbeth model, *FamMac*[4]. It is used besides others by Chordia *et al.* (2001), since this type of model has the most observations. In these models excess returns are used as dependent variable. The *FamMac* model I define as follows:

$$\begin{aligned} Ret = \alpha_0 + \alpha_1 Vol + \alpha_2 Size + \alpha_3 BM + \alpha_4 Price + \\ + \alpha_5 Ret_{2-3} + \alpha_6 Ret_{4-6} + e, e \sim N(0, \sigma^2), \end{aligned} \tag{3.7}$$

where *Ret* is excess return, *Vol* is trading volume, *Size* is the natural logarithm of the firm's market value of the equity, *BM* is the natural logarithm of the book value of equity to market value of equity. Moreover, Ret_{2-3} and Ret_{4-6} record the returns in previous periods. If one of the variables *Size* and *Price* is missing, I classify the model as *FamMacB* type of model. This one is used, for instance, by Lewellen (2015)

[4]The name of any variable is always written in italics with capital letter.

or Brandle (2010). Both mentioned Fama - Macbeth models employed panel data with longer time periods (usually gathered monthly). Contrary to Fama - MacBeth models stands VAR[5] models which use time series data with higher frequency (daily frequency of data). I split these models into three groups according to the number of lags included in their base equations. In *Vara* type of model, I include estimates from VAR equations, which includes one lag from both variables at the maximum:

$$\begin{aligned} Ret_t &= \alpha_0 + \alpha_1 Vol_t + \alpha_2 Vol_{t-1} + \alpha_3 Ret_{t-1} + e_t, e_t \sim N(0,\sigma^2), \\ Vol_t &= \beta_0 + \beta_1 Ret_t + \beta_2 Ret_{t-1} + \beta_3 Vol_{t-1} + u_t, u_t \sim N(0,\sigma^2). \end{aligned} \tag{3.8}$$

This specification is used, for example by Rotila *et al.* (2015) or Yonis (2014).

The second group of VAR models (*Varb*) maintain the structure of the first equation in Equation 3.8, but in the second equation replaces lag of returns by second lag of volume. I can see this specification in researches written by Lee & Rui (2002) and Ciner (2002) among others. The VAR equation is adjusted in following form:

$$\begin{aligned} Ret_t &= \alpha_0 + \alpha_1 Vol_t + \alpha_2 Vol_{t-1} + \alpha_3 Ret_{t-1} + e_t, e_t \sim N(0,\sigma^2), \\ Vol_t &= \beta_0 + \beta_1 Ret_t + \beta_2 Vol_{t-1} + \beta_3 Vol_{t-2} + u_t, u_t \sim N(0,\sigma^2). \end{aligned} \tag{3.9}$$

The last VAR specification (*Varc*) is for those VAR models with two (Saatcioglu & Starks 1998), three (Louhichi 2012) or five (Pisedtasalasai & Gunasekarage 2007) lags for both variables in both equations in the models. Moreover, time series in all VAR models are tested for a unit root by Augmented Dickey - Fuller test or by Phillips - Perron regression or by both in primary studies. For more detailed description of these tests follow Lee & Rui 2002. Furthermore, volume time series is usually detrended.

Next two groups of models originate from the basic equation:

$$Ret_t = \alpha_0 + \alpha_1 Vol_t + e_t, e_t \sim N(0,\sigma^2), \tag{3.10}$$

where sometimes Vol_{t-1} is used. The first specification employed by Shu *et al.* (2004) or Tapa & Hussein (2016) among others estimate model mentioned in Equation 3.10 simply by OLS. Therefore, this specification is called *Simple*. Other authors as Sana Hsieh (2014) and Tahir *et al.* (2016) improve variance equation by GARCH[6] in order to capture heteroskedasticity:

[5] Vector Autoregression.
[6] Generalized autoregressive conditional heteroskedasticity.

$$\begin{aligned} Ret_t &= \alpha_0 + \alpha_1 Vol_t + e_t, \\ e_t|(e_{t-1}, e_{t-2} &\sim N(0,h)|, \\ h_t = \sigma_t^2 &= \beta_0 + \beta_1 e_{t-1}^2 + \beta_2 h_{t-1}. \end{aligned} \tag{3.11}$$

To this group, called *Garch*, I add those observations, that estimate variance equation by some kind of Egarch[7]. Such studies are, e. g. Mahajan & Singh (2009a) and Kim (2005).

Finally, the last group of models is named *OthMod*, since these models do not fit to any of nine above mentioned groups. Fundamentally, these are threshold models (Hafner 2005), or models with larger number of lags (Vo 2017).

3.2.6 Other dummies and estimation methodology

Besides the basic models I describe in the previous paragraph, some authors add to their models dummy variables to capture differences between particular observations. Among these variables belongs Monday, or January effects, market beta, portfolio beta or sales-to-price ratio. I consider these dissimilarities and create dummy variables for each such difference, which is recognized in more than 2% cases (i. e. more than 11 single observations). All these variables are listed in an overview at the and of this chapter.

The methodology of estimation differs among articles as well. Most of them employ *Ols* or *Gmm*. *Ols* is used, for instance, by Tapa & Hussein (2016), or by Narayan & Zheng (2010). On the contrary, Lee & Rui (2002) and Ciner (2003) among others prefer *Gmm*. Few authors use *Mle* (Epps & Epps 1976). If the method is not clear or specified, I capture it in for other estimation methods (*Othest*).

3.3 Value-added Variables

Among value-added variables I count those, which are not available during performing the primary study (Stanley & Doucouliagos 2012). These pieces of information are relevant and usually "study-invariant". Its relevance consists in ability to explain variation from different studies. I find inspiration in collecting such variables in Bajzik (2017) and in Astakhov *et al.* (2017). I am interested in whether the study is published or not (captured by variable *Pblshd*). Next I investigate in how influential

[7]Exponential generalized autoregressive conditional heteroskedasticity.

journal the paper is published by adding variable *Impact*[8]. This variable is downloaded from a discounted recursive impact factor from RePEc Ideas[9]. There are many ways to set the impact factors. I choose the one from RePEc since it reflects the quality of citations and includes almost all economic journals, even the working papers series. For previous use of the recursive impact factor in meta-analysis framework refer, for example, to Valickova *et al.* (2015), or to Rusnak *et al.* (2013). Besides I am curious about the impact of the study itself. Therefore, I include the number of citations (variable *Cit*) from Google Scholar[10]. Moreover, these three variables are useful in identification of potential publication bias, which even augment their relevance for the study.

Finally, I add three variables for capturing geographical differences. I add variable *Msize*, which stands for market size and helps me to distinguish bigger markets from smaller ones. I collect these pieces of information from World Bank[11] in terms of GDP in billions of US dollars in midyear of the data. Only information for Taiwan I need to collect from National Statistic Republic of China[12], since World Bank does not provide information for Taiwan. Because I get these values in terms of New Taiwanese dollars, I recalculate it to US dollars by midyear of data NTD-USD fx-rate according to Federal Reserve Bank[13]. For each midyear I use end year fx-rate. Just for year 1981 I use the fx-rate from 31^{st} December 1983, since former data are not available. Otherwise, I collect information about a region, captured by variable *Region* and if the evidence was for US I want to distinguish between different stock exchanges. Finally, I have dropped from this idea, since some of the studies (e. g. Chordia *et al.* 2001, Brennan *et al.* 1998) use data from more than one US Stock Exchange (NYSE + AMEX), others (e. g. Lee & Rui 2000, Kim 2005) prefer to use indexes (S&P500), which I already distinguish in the dataset. Therefore, I do not difference these in the dataset more, since the necessary dissimilarities are already captured.

[8]This variable is without a unit.

[9]RePEc IDEAS: Recursive Discounted Impact Factors for Series and Journals [online]. Infogram: ©1997 [cit. 12. 2. 2019]. Available from: https://ideas.repec.org/top/top.series.rdiscount.html.

[10]Google Scholar [online]. Infogram: ©2004 [cit. 12. 2. 2019]. Available from: https://scholar.google.cz/.

[11]The World Bank: GDP [online]. Infogram: ©2016 [cit. 18. 2. 2019]. Available from: hhttps://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

[12]National Statistic Republic of China [online]. Infogram: @2019 [cit. 18. 2. 2019]. Available from:https://eng.stat.gov.tw/point.asp?index=1.

[13]Federal Reserve Bank: Real Effective Exchange Rates [online]. Infogram: ©2016 [cit. 18. 2. 2019]. Available from: https://fred.stlouisfed.org/graph/?id=AEXTAUS.

3.4 Final Adjustments

When I collect the data, there is need of some adjustments. At first, I drop one dummy variable from each group of dummies. That means, for instance, from the group of type of data - panel, time series, cross-sectional, I drop one. I usually get rid of the one with the most observation, since the dropped variables become the benchmark in interpretation of the result.

Moreover, I transform three variables into logarithms - *Total* number of observations, number of *Citations*, *Market size* - and I adjust the *Midyear* of data and the publication year of the study (*Pubyear*) by subtracting 1965. All these transformations are visualized in Appendix A in figures Figure A.2 to Figure A.5.

In addition to, I employ winsorizing to contend with outliers in dependent variable and in relative standard errors. The winsorizing sets the values of the lowest and highest estimates to desired percentiles. I set the highest and the lowest estimates to 99.5% percentile and 0.5% percentile respectively, since the gathered dataset is rather small and homogeneous.

Finally, to guarantee that each paper has the same weight I multiplicate the data by the inverse number of the estimates per study. It is because the number of estimates vary across articles. For example, Mahajan & Singh (2008) provide one result. On the contrary, from Hu (1997) I draw 62 different estimates.

3.5 Overview of used variables

The overall overview of variables currently in use is to be found in following table. I drop one variable from each dummy variable trap. There are two results at *Retvol* and *Se*. The ones without brackets are before winsorizing at 1% level and these in brackets are those after this adjustment. The negligible differences between non-winsorized and winsorized estimates shows presence of small number of outliers. The median values naturally remain the same. Besides the difference between the median and mean of the return-volume relationship detects possible publication bias. *Total* number of observations, *Citations* and *Market size* are in logarithms. *Midyear* is adjusted by -1965.

Table 3.1: Variables in Use

Label	*Description*	*Mean*	*SD*	*Med*
Retvol	Estimates of return-volume relationship (explained variable)	0.02 (0.02)	0.17 (0.16)	0.00 (0.00)
Se	Estimates of standard errors of return-volume relationship	0.04 (0.04)	0.06 (0.06)	0.02 (0.02)
Contem	=1 if the return-volume relationship is contemporaneous	0.61	0.49	1.00
Spillover	=1 if the relationship indicates spillover from one market to another	0.09	0.29	0.00
Absolute	=1 if the returns are in absolute terms	0.11	0.31	0.00
Abnormal	=1 if the returns are defined as abnormal	0.02	0.15	0.00
Excess	=1 if the returns are defined as excess	0.24	0.43	0.00
Volume	=1 if the volume is expressed in terms of dollar volume of the trade	0.12	0.32	0.00
Shares	=1 if the volume is expressed in terms of shares traded	0.39	0.49	0.00
Detrend	=1 if the volume series was detrended	0.19	0.39	0.00
Asia	= 1 if the observation is linked to Asia	0.39	0.49	0.00
Europe	= 1 if the observation is linked to Europe	0.12	0.32	0.00
Australia	= 1 if the observation is linked to Australia	0.04	0.21	0.00
Index	=1 if the cumulative returns value for stocks from particular index was used	0.33	0.47	0.00
Bank	=1 if the research relates only to banking sector	0.03	0.16	0.00
Nonfin	=1 if the research relates only to some non-financial sector	0.08	0.27	0.00
Othinv	=1 if the research relates to other specific market	0.13	0.34	0.00
Time	=1 if the time series data were used	0.44	0.50	0.00
Cross	=1 if the cross-sectional data were used	0.01	0.12	0.00
Hourly	=1 if the data were collected hourly or more frequently	0.10	0.30	0.00
Daily	=1 if the data were collected daily	0.33	0.47	0.00
Weekly	=1 if the data were collected weekly	0.06	0.24	0.00
Length	Length of time period	13.85	10.47	10.00
Midyear	Midyear of data	26.23	12.01	26.50
Deving	=1 if the estimate is for developing country	0.25	0.44	0.00

Continuation of Table 3.1				
Label	*Description*	*Mean*	*SD*	*Med*
Total	Total of observation	8.51	3.11	7.87
Msize	Market size in terms of GDP (billions of dolars) in midyear of data (in logarithms)	6.97	1.46	7.06
FamMac	=1 if first specification of the Fama - Macbeth model is used	0.20	0.40	0.00
Vara	=1 if the first specification of VAR model is used	0.09	0.29	0.00
Varb	=1 if the second specification of VAR model is used	0.06	0.24	0.00
Varc	=1 if the third specification of VAR model is used	0.06	0.24	0.00
Simple	=1 if the simle model estimated by OLS is used	0.28	0.45	0.00
Garch	=1 if the ARIMA with GARCH in error term is used	0.15	0.35	0.00
OthMod	=1 if other than specified above model specification is employed	0.02	0.15	0.00
Monday	=1 if effect of Monday or January trading is considered	0.02	0.16	0.00
Yld	=1 if the dividend yield as measured by the sum of all dividends paid over the previous 12 months, divided by the share price at the end of the second to last month is incorporated	0.13	0.33	0.00
Stdev	=1 if some measure of standard deviation is added in the model	0.13	0.33	0.00
MarBet	=1 if variable represents market beta is included	0.06	0.24	0.00
Illiq	=1 if the average ratio of the absolute daily stock returns to its dollar trading volume is included	0.07	0.25	0.00
Accrual	=1 indicates inclusion of variable measuring change in non-cash net working capital minus depreciation in the prior fiscal year	0.02	0.14	0.00
SaleP	=1 if sales to price ratio is added	0.03	0.18	0.00
FirmBet	=1 if firm or portfolio beta is included in the model	0.07	0.25	0.00

Continuation of Table 3.1				
Label	*Description*	*Mean*	*SD*	*Med*
Size	=1 if natural logarithm of firm's market capitalization is included	0.02	0.16	0.00
Trimmed	=1 if the primary dataset was trimmed	0.08	0.28	0.00
ExclJan	=1 if all months but January are included in the primary dataset	0.07	0.26	0.00
Informed	=1 if measure of the probability of information-based trading in the previous year is included	0.03	0.17	0.00
Ivol	=1 if idiosincratic volatility is explained variable in primary study	0.03	0.16	0.00
Mle	= 1 if MLE estimation method was used	0.04	0.19	0.00
Gmm	=1 if GMM estimation method was used	0.35	0.48	0.00
Othest	=1 if other types of estimation were used	0.21	0.41	0.00
Impact	Discounted recursive impact factor from RePEc IDEAS	0.59	0.93	0.06
Cit	Number of citations (in logarithms)	4.21	2.38	4.48
Pblshd	=1 if the article was published	0.81	0.40	1.00

Notes: This table shows mean, standard deviation and median for each variable used in own estimation. The average partial effect method (Wooldridge 2015) was used for means of all variables in logarithms.

Chapter 4

Methodology

4.1 Publication Selection Bias

During the analysis I need to contend with publication selection bias. It is a well-known phenomenon and the Meta-Regression Analysis (MRA) is very convenient tool for detecting it. The danger of publication bias lie in affecting weights of estimates that are used in meta-analysis. Thus, I test for the publication selection before I move forward to analysis of heterogeneity. The publication selection bias arises due to different probability of reporting of miscellaneous estimates. The discrimination is based mainly on unintuitive sign, statistical significance or on the magnitude of the estimate. Researchers may have different motivation for such a hiding. For instance, they do not want to go against mainstream results, because the mainstream results are easier to publish (Havranek & Irsova 2015).

The publication selection is not widely spread only in economics, but mainly in medical science. Therefore, now the best medical journals require registration of all clinical trials prior publication of articles. It means that even though some results are not published, the researchers can find the results for all trials. Similarly to medical science, the American Economic Association has decided to establish a register of randomized experiments in order "to counter publication bias" (Siegfried 2012).

About persistence of this phenomenon in economics write, for instance, Doucouliagos & Stanley (2013). In their research, they find out that most fields of empirical economics are affected seriously by publication selection bias. For example, Astakhov *et al.* (2017) discover that the researchers in area of firm size and stock returns prefer to report estimates that show a negative relation between size and returns and that are statistically significant. It exaggerates the mean reported coefficient three times. Furthermore, they show that the publication bias has been decreasing over time. This conclusion is not surprising, since after emphasizing phenomena of publication bias, authors naturally started reporting more precisely. On the other hand, Havranek & Irsova (2015) find no evidence for selective reporting in the field of border effect.

4.1.1 Funnel plot

In analysis of publication bias, I start with funnel plot. It is easiest way to detect publication bias proposed by Egger *et al.* (1997). On horizontal axis of this scatter plot is depicted the magnitude of the estimated effects and on the vertical axis is shown the precision measured by the inverse of the estimated standard error. The precision of the estimates depends on their distance to the mean underlying effect. It is in case that the research is not affected by publication selection. There are two possible issues that can arise from the observation of the scatter plot. Firstly, the estimates can get more spread shaping a symmetrical reversed tunnel. It happens when precision decreases. Secondly, in presence of publication bias, the funnel plot is asymmetrical (estimates of a particular sign or magnitude are incorrectly rejected by the researchers), or it becomes hollow (statistically insignificant estimates are discarded). In the worst case, the scatter plot can be both asymmetrical and hollow Egger *et al.* (1997). The results of funnel plot related to this study are depicted in Section 5.1.

4.1.2 Formal tests

I employ more formal tests of publication bias, since the funnel plot is only a simple visual tool for publication selection evaluation. I follow the methodology of Stanley & Doucouliagos (2012), since they proposed the baseline for the study of the publication bias. Moreover, I use some of the latest innovations proposed by Astakhov *et al.* (2017). The following base regression clarifies the relationship between observed effect (Expected stock returns) and its standard error ($SE(r_{ij})$).

$$r_{ij} = \beta_0 + \beta_1 SE(r_{ij}) + e_{ij}, e_{ij} \sim N(0, \sigma^2), \tag{4.1}$$

where r_{ij} is the i-th estimate of the partial correlation coefficient between expected stock returns and trading volume from study j.

In case of no publication bias presents, β should be zero, since the ratio of the reported estimates of the size effect to their standard errors follows a t-distribution. The econometric techniques used by the researchers to estimate the size effect guarantee this property. On the contrary, if particular sign of the size effect is preferred, it leads to non-zero β due to heteroskedasticity of the regression. In similar manner, if there is preference for reporting statistically significant estimates, large point estimates need to be reported to counterweight standard errors. It again leads to a non-zero β. Hence, the hypothesis $H_0 : \beta = 0$ arises. It is also known as funnel-asymmetry test.

I estimate Equation 4.1 by five different estimation methodologies. First, I use simple OLS with standard errors clustered at the level of individual studies. The clustering corrects for possible heteroskedasticity. Second, I run panel data regression employing between effects. Third I follow Stanley & Doucouliagos (2012) and Bajzik (2017) in multiplying the regression Equation 4.1 by $1/SE(r_{ij})$. This estimation procedure is known as weighted least squares (WLS) and assigns more weight to more precise studies and directly contend with heteroskedasticity. Therefore the weight $1/SE(r_{ij})$ is called *Precision*. From this modification of Equation 4.1 stems the following regression:

$$\frac{r_{ij}}{SE(r_{ij})} = \frac{\beta_0}{SE(r_{ij})} + \beta_1 + u_{ij}, u_{ij} \sim N(0, \sigma^2), \tag{4.2}$$

where the explained variable is reported t-statistic and the selective reporting bias is measured by β.

In fourth specification I use instead of *Precision* the inverse number of estimates per study as a weight. Fifth, since there is a problem of endogeneity in Equation 4.2 that means some omitted study characteristics, for instance, estimation techniques that influences both estimates and their standard errors I choose instrumental variable (IV) regression as proposed in Zigraiova & Havranek (2016) or in Astakhov *et al.* (2017) as a third approach for detecting publications selection bias. Such an instrument can be correlated with the standard error of the estimated coefficient (relevance condition), but it is not correlated with the choice of the estimation technique (exogeneity condition). Thus, they estimate the IV regression using the $1/\sqrt{n}$ as an instrument (inverse of the square root of the number of observations), which obviously meets both condition and it is thus valid instrument for IV estimation. For example, Havranek & Irsova (2012) make a discovery that studies with small datasets tend to engage more in publication selection.

Next I test publication selection bias by inclusion of interaction terms between standard error with the year of publication and between standard error with the recursive *Impact* factor (reported on the IDEAS/RePEc website) to the Equation 4.1. These interactive terms are proposed by Astakhov *et al.* (2017). The effect of a study's publication year on the strength of the selective reporting bias is not certain. But based on the development in area of econometric techniques in the last years might bring results closer to the true effect. Moreover, since the question of publication selection is more discussed in the last years, there is systematic pressure on publishing all results, not only those that are expected. Based on these two observations, the hypothesis is that the extent of publication bias decreases in time. It would be in line, for example, with economics-research-cycle hypothesis. This conclusion is confirmed by Astakhov *et al.* (2017) in the field of firm size and stock premium, which is close area to ours. The same is found by Havranek & Irsova (2012) in area of foreign direct investments. Similarly, the uncertainty is present regarding the effect of a study's publication in more renown journals. Researchers might be hesitant to submit studies with "unexpected" or inconclusive results because of fear of rejection. On the other hand, higher quality journals uses more strict review procedures. Findings of Astakhov *et al.* (2017) are inconclusive regarding this phenomenon. On the other hand, Havranek & Irsova (2012) found that publication bias is obvious among studies published in peer-reviewed journals.

4.1.3 Advanced techniques

Besides commonly used and widely known publication bias detection techniques I employ three advanced techniques for publication bias investigation that was developed recently. They arose in the economic literature since the public often question the relevance and reliability of the models described in the section Subsection 4.1.2. Estimating β_0 from Equation 4.1 gains an unbiased estimate of the mean cleansed for publication bias only if publication selection is proportional to its standard error. However, in practice I deal with unknown functional form of the publication selection procedure. Therefore, the first advanced technique I employed is the advanced estimator introduced by Andrews & Maxmilian (2019). Their estimator is to my knowledge the only one who addresses the detected problem and is probably unbiased under any form of publication selection as it is proposed in Havranek & Sokolova (2019). On the other hand, it is unclear whether and how the heterogeneity affects the properties of this estimator or not. Besides as I can check, for instance, in Stanley & Doucouliagos (2014) the Equation 4.1 has been examined thoroughly by several Monte Carlo studies. Moreover, it has been found that this specification works well especially in case the corrected effect is small, which is even case of my dataset. Besides the basic specification allows to test the exogeneity of the standard error as I did by IV estimator.

Next I employed method introduced by Stanley *et al.* (2017), the weighted average of the adequately powered (WAAP). It is an alternative to the random-effects estimator. As an indicator of statistical power serves the precision. The larger the precision the higher the power of the connected study to detect the effect. Stanley *et al.* (2017) used the Cohen (1988)'s standard of 80% to define "adequate" power. WAAP is an unrestricted weighted average upon only those adequately powered estimates. The base regression for computation is the one I use in *Precision* specification - Equation 4.2 (without constant). In this case only the m adequately powered estimates are employed instead of all. These are defined based on results standard error from the *Precision* equation. The adequately powered estimates are those, whose standard errors are smaller than $|SE/2.8|$, where SE is the standard error from the results of the base equation. The number 2.8 is a sum of 1.96, which serves for testing the significance of the null hypothesis at 5% level and 0.84, which is traditional conventional standard for statistical power and significance. For more information refer to Stanley *et al.* (2017). Since the dataset is quite small, I decide to use constant 2.485 instead of 2.8 as a sum od 1.645 (testing the null hypothesis at 10% significance level) and 0.84. It helps me to get sufficient number of observation to perform the computation. The weakness of this approach lies just in likewise situation, when low number of adequately powered studies is in the research area. This concession may indicate that the observed true effect is not significant. In such cases, usually the WAAP cannot be computed, and the hybrid between WAAP and *Precision* (WLS) is used. On the contrary, WAAP approach dominates in situations with numerous high and low powered studies.

As the last but not least method to study publication bias in the field of the relationship between expected stock returns and trading volume I propose the methodology recently released by Furukawa (2019). This approach is commonly known as stem-based method. As the name indicates, similarly to WAAP, it is based on the most precise estimates. This stem-based method alleviates publication bias in a way that is robust under various assumptions. The method is fully data-dependent and non-parametric. Generally, the results originating from this approach are more conservative than from other commonly used methods. On the other hand, one shortcoming originates in the assumption that the most precise study reasonably approximates the true mean (on average). Besides it offers a formal criterion to choose the optimal number of most precise studies to conduct the meta-analysis estimation. The baseline equation for choosing the right number of estimates (n) proposed by Furukawa (2019) is the Mean Square Error (MSE) optimization of the following equation:

$$\min_{n} MSE(n) = Bias^2(n) + Var(n), \tag{4.3}$$

where the bias increases with number of the observation n, since less precise estimates will be included, while the variance decreases in n. For more details of this methodology, please refer to Furukawa (2019), as it is mentioned above .

4.2 Bayesian Model Averaging

Besides publication selection investigation my main focus is on determining the most influential factor in expected stock returns estimation. I consider following regression:

$$r_{ij} = \alpha_0 + \sum_{k=1}^{52} \beta_k X_{k,ij} + e_{ij}, \tag{4.4}$$

where $X_{k,ij}$ labels the value of a k-th explanatory variable for an i-th estimate from a j-th study. Since I believe that every variable can contribute to explaining the heterogeneity among the estimates, I employ the Bayesian Model Averaging (BMA) for this inquiry. It is classical method in the field of meta-analysis. BMA was pioneered by Raftery *et al.* (1997) and Raftery (1995) in social sciences. In economics it found stable place in determining of economic growth (Fernandez *et al.* 2001, or Durlauf *et al.* 2008). Lately, BMA has attention more substantially across different economics areas. A brief survey about usage BMA in economics is conducted by Moral-Benito (2011). As the name of the method suggests BMA estimation procedure is not based on fitting the best model, but it uses weighted average of all possible combinations of linear models provided from the data set. It is very beneficial in case of uncertainty about the specification of the regression model, when there is several competing theories all propose different regression model. In other words, the advantage of BMA methodology against model selection approaches is grounded in jointly testing the relevance and importance of various concepts and theories. Moreover, it addresses omitted variable bias in systematic manner. Furthermore, the BMA methodology is thoroughly robust to outliers, because it assigns lower weights to their individual regression, because of their lower fit (Horvath *et al.* 2017). For the following comprehensive description of BMA methodology, I follow Horvath *et al.* (2017), if it is not stated differently.

4.2.1 Foundation

For illustration of BMA I consider the following linear model:

$$y = \alpha + \beta X + e, e \sim N(0, \sigma^2 I), \tag{4.5}$$

where y is the explained variable, α is a constant, X is the matrix of independent variables, β are their corresponding coefficients, and e represents a vector of normally distributed iid[1] error terms with variance σ^2.

[1]Independent and Identically Distributed.

In BMA are considered all possible combinations of X from Equation 4.5. BMA takes the weighted average of their coefficients (more concretely, this weighted average is based on a subset of the models using MCMC sampling, which is discussed below). The structure of the subset of the models can be expressed as follows:

$$y = \alpha_i + \beta_i X_i + e, e \sim N(0, \sigma^2 I), \tag{4.6}$$

where X_i is a subset of X and the corresponding coefficients are α_i and β_i. When assume that there is K possible explanatory variables, the total number of possible models equals to 2^K and $i \epsilon [1, 2^K]$.

It is derived from Bayes' rule that

$$p(\beta|y, X) = \frac{p(y, X|\beta)p(\beta)}{p(y|X)}. \tag{4.7}$$

Here, $p(\beta|y, X)$ stands for the posterior density, $p(y, X|\beta$ represents the marginal likelihood (ML, it is known as data-generating process), $p(\beta)$ is the prior density, and $p(y|X)$ expresses the probability of the data. As mentioned above, in BMA is compared numerous different models $M_1, ..., M_i$, where $i \epsilon [1, 2^K]$, with K possible regressors. M_i depends on the parameters β_i. Their posterior probability can be depicted as follows:

$$p(\beta|M_i, y, X) = \frac{p(y|\beta_i, M_i, X)p(\beta_i!M_i)}{p(y|M_i, X)}. \tag{4.8}$$

Now I move to description of the individual components of Equation 4.7 and to the averaging principle of the BMA.

4.2.2 Posterior Model Probability

One of the fundamentals in BMA framework is the posterior model probability (PMP). It takes the submodels and assigns them weights for averaging model. It also emerges from Bayes' theorem:

$$\begin{aligned} p(M_i|y, X) &= \frac{p(y|M_i, X)p(M_i)}{p(y|X)} \\ &= \frac{p(y|M_i, X)p(M_i)}{\sum_{s=1}^{2^k} p(y|M_s, X)p(M_s)}, \end{aligned} \tag{4.9}$$

where *p(y|X)* is the integrated likelihood. The integrated likelihood is constant over all models and hence it is interpreted as a multicaptive term. The *(*p(y|M_i) is the marginal likelihood (ML) of the model. The ML is the probability of data given by the model M_i. And the last term in the Equation 4.9 is the prior model probability $p(M_i)$. Thus, the PMP is proportional to the prior probability and the ML. It is very common to set the prior probability to uniform one ($p(M_i) \propto 1$). It reflects the lack of knowledge about the true model. Reader may find it, for instance, in Havranek & Irsova (2015). The model prior is set by the researcher and it shows the initial beliefs before scrutinizing the data.

Now there remains just one step to the posterior inclusion probability (PIP), which is usually reported in the standard BMA framework. It reflects the probability that a particular explanatory variable is included in the "true" model. PIP is the sum of the PMPs of all the models that includes the explanatory variable k in the following equation:

$$PIP = p(\beta_k \neq 0|y, X) = \sum_{i=1}^{2^k} p(M_i|\beta_k \neq 0, y, X). \tag{4.10}$$

4.2.3 Posterior Mean

The focus of research are often the point estimates and it is possible to derive them within the Bayesian framework. Moral-Benito (2011) suggests that the weighted posterior distribution of any statistic (β coefficients - Equation 4.5) is gained from the following formula:

$$p(\beta|y, X) = \sum_{i=1}^{2^k} p(\beta|y, X, M_i)p(M_i|y, X), \tag{4.11}$$

where $p(M_i|y, X)$ is the PMP of the particular model M_i from the Equation 4.9. By taking expectations across the following equation the desirable point estimates are obtained:

$$E(\beta|y, X) = \sum_{i=1}^{2^k} E(\beta|y, X, M_i)p(M_i|y, X), \tag{4.12}$$

where $E(\beta|y, X)$ stands for the averaged coefficient, and $E(\beta|y, X, M_i)$ is the estimate of the coefficients β_i from model M_i. Besides, the posterior distribution of the coefficients depends on the choice of the prior g. The expected value of the parameter in M_i can be expressed as follows:

$$E(\beta|y, X, g, M_i) = \frac{g}{1+g}\hat{\beta}_i, \tag{4.13}$$

where $\hat{\beta}_i$ is the estimate from standard OLS.

4.2.4 Model Priors and Parameter Priors

There is a requirement for two types of priors in the BMA framework: $p(M_i$ on the model space and g on the parameter space. The priors are seminal in determining posterior parameter space (e. g. Ciccone & Jarocinski 2010). Now I discuss both of the priors briefly.

One of the most well-known model priors is the uniform model prior. It arises from the binomial distribution, where each of the independent variables is included in the model with probability of success θ. In this case, the prior probability of model M_i with k_i independent variables given θ is:

$$p(M_i) = \theta^{k_i}(1-\theta)^{K-k}. \tag{4.14}$$

To achieve the uniform distribution, it is sufficient to set $\theta = \frac{1}{2}$. It assigns equal probability $p(M_i) = 2^{-K}$ to all models.

Despite the fact that the uniform model prior tends to assign greater weight to intermediate model sizes, I employ it in this research.

As I already mentioned, I use Zellner's g prior structure. It is common approach in the literature, when one assumes that the priors on the constant and error variance from Equation 4.6 are evenly distributed. It means that $p(\alpha_i) \propto 1$ and $p(\sigma) \propto \sigma^{-1}$. In my research it is assumed that β_i coefficients follow the normal distribution. In addition to, I need to formulate beliefs regarding their mean and variance before data investigation. A conservative approach, which is very spread among researchers assume a conservative mean of 0. It reflects lack of prior knowledge with respect to coefficients. The variance structure of Zellner's g is defined as $\sigma^2(g(X_i'X_i)^{-1}$. From these two assumptions I can derive the structure of coefficient dependence on the prior g as:

$$\beta_i|g \sim N(0, \sigma^2(g(X_i'X_i)^{-1}). \tag{4.15}$$

In other words, the posterior variance of drawing from the sample is proportional to the prior variance of the coefficients. The balance between prior variance and posterior variance from the data is set by the parameter g (Feldkircher & Zeugner 2009). Small g leads to low variance in the prior coefficients. Thus, the coefficients are reduced to zero. On the contrary, a large g assigns higher weight to the data.

Among the most popular choices of the parameter g belongs Unit Information Prior (UIP). It sets $g = N$. Next one is Benchmark Risk Inflation Criterion (BRIC), which set $g = max\{N, K^2\}$. These two are known as "fixed-g" priors, since the parameter prior is set for all the possible model. Different it is with the third popular choice of g, so-called hyper-g. It allows updating the prior for individual models. Moreover, it penalizes including new variables into the model. The new covariates are regulated through inclusion of hyperparameter g in the marginal likelihood (Feldkircher *et al.* 2014)[2]. Finally, I decide to use UIP, even though hyper-g prior should lead to more stable posterior structure. For use of hyper-g prior instead of "fixed-g" priors I will wait for more researches, which confirms its more stable posterior structure.

4.2.5 MCMC Sampling

MCMC sampling is used because of computational difficulty of BMA when there is large number of potential independent variables K. Because with K independent variables there is 2^K possible models, it is technically impossible to estimate all of these models. I can even say it is infeasible. In case there is less than 15 potential explanatory variables I can proceed to BMA estimation with all possible models, otherwise I use MCMC sampling. The MCMC samplers provide sufficient approximation for the crucial part of the posterior model distribution. It contains the most likely models. The chain use in MCMC sampler applies the Metropolis-Hastings algorithm. The logic behind can be described as follows (Zeugner 2011):

The sampler stand at current model M_i with PMP $p(M_i|y,X)$ at any step i. In the step $i+1$ different model M_j is suggested to replace current model M_i. The new model is accepted with the following probability:

$$p_{i,j} = min\left(1, \frac{p(M_j|y,X)}{p(M_i|y,X)}\right). \tag{4.16}$$

If the model M_j is accepted, it becomes the current model and it is compared to new model M_k according to two steps described above. In case the model M_j is rejected, the step $i+2$ takes place. In this step the model M_k is proposed to challenge the model M_i. With increasing number of the iterations, the number of times that each model is kept converges to the distribution of PMP $p(M_i|y,X)$.

There are two basic MCMC samplers for drawing models, the Birth-death sampler and the Reversible-jump sampler. The Birth-death sampler chooses randomly one covariate which is included in the data and adds it to the current M_i model, or it drops it from current M_i model, if it is already included in it. The Reversible-jump sampler with 50% probability randomly swaps one of the explanatory variables in the current model M_i for an explanatory variable previously excluded from M_i. And with 50% probability, the Birth-death sampler is employed to determine the next candidate model within the Reversible-jump sampler.

It is possible that the MCMC sampler begins with a model, which is not "good" one (that means it has low PMP). Therefore, there is a predefined number of initial

[2]See Liang *et al.* (2008) for more details about hyper-g prior.

draws, so-called burn-ins, which are usually removed. Next, the correlation between the PMP derived from the MCMC sampler and the PMP derived from an analytical approach is used as a measure of the quality of the approximation of the sampler. It depends directly on the likelihood of the initially selected model and on the number of draws (iterations). The correlation above 0.9 indicates a "good degree of convergence" (Zeugner 2011), otherwise the number of sampler iterations should be increased.

4.3 Frequentist Model Averaging

Moreover, I employ Frequentist Model Averaging (FMA) in the robustness check. It is an alternative approach to BMA. I follow Havranek *et al.* (2017), which is probably first meta-analysis that use FMA. It utilizes for individual regressions the standard technique of the literature on estimated dependent variable models. The intuition behind FMA is similar to BMA, which is discussed in depth in the previous section. There are many models that combines different subsets of explanatory variables. Then they are estimated and weighted based on their goodness of fit and parsimony.

The disadvantage in comparison to BMA is computational difficulties of FMA in this area. The few studies that depend only on frequentist techniques usually use information criteria as weights. These are for instance AIC or BIC. Nonetheless, Hansen (2007) shows that asymptotically optimal are weights selected by minimizing the Mallows criterion. In a nutshell, the Mallows criterion takes from the model average fit an estimate of the average squared error.

Next problem to address is simplification of the model space in FMA. It is unfeasible to estimate all 2^K models, which is 2^{52} in my case. Moreover, I cannot employ MCMC algorithm I use in BMA. Therefore, I use orthogonalization of the variable space and thus reduce the number of models need to be estimated from 2^{52} to 52. The inverse-variance weights are used in individual regressions to account for the estimated explained variable matter. For more details see Amini & Parmeter (2012) and Magnus *et al.* (2010).

Chapter 5

Results

5.1 Publication Bias

As is proposed in Subsection 4.1.1 I begin the analysis of publication bias with visual investigation of the funnel plot suggested by Egger *et al.* (1997). On the y-axis is captured the precision of the size coefficient and the point estimates of the size coefficient are then depicted on the x-axis.

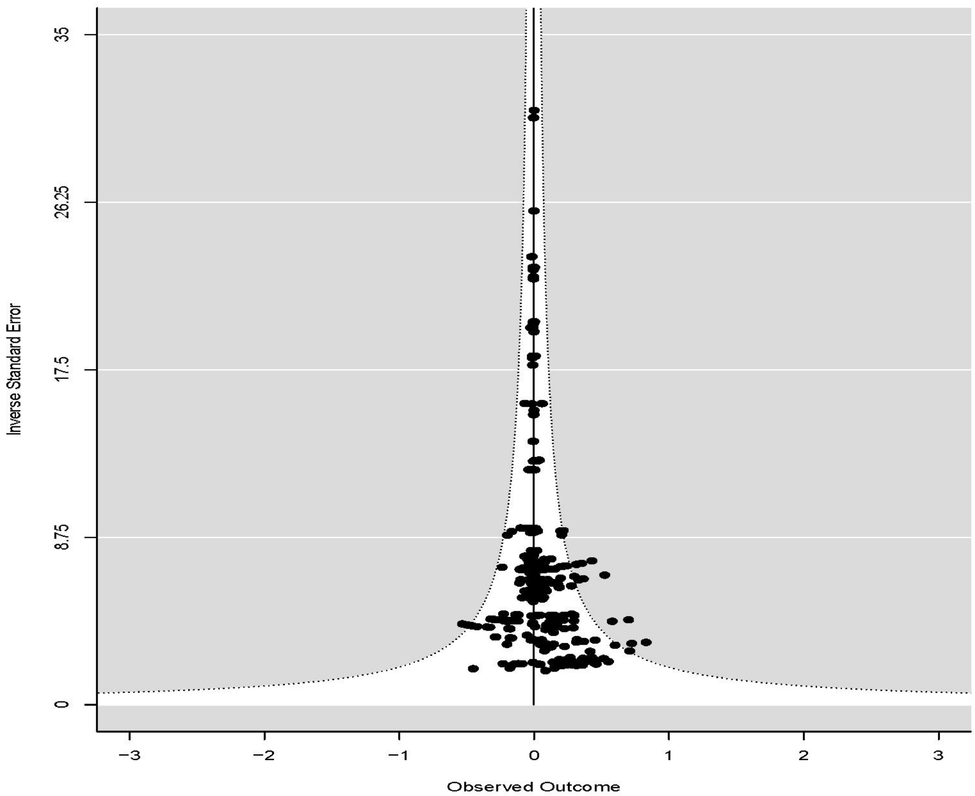

Figure 5.1: Histogram of Inverse Standard Errors

When there is no selective reporting bias, the funnel plot takes symmetrical shape with the most precise estimates gathered around the underlying mean value of the size effect. Whereas less precise estimates are scattered around the mean. The Figure 5.1 presents evidence of slight selective reporting, since the plot is asymmetric with point estimates gathered in the right-hand tail. Hence, other things being equal, the probability of reporting negative estimates of the relationship between expected stock returns in academic studies is less than in case of positive estimates.

I proceed by testing for publication bias in the formal way using Equation 4.1 and Equation 4.2 from the section Subsection 4.1.2. In the Table 5.1 in the Column (1) is the baseline result of OLS estimation the partial correlation coefficient on its standard error. The β_1 coefficient is both positive and significant, which indicates a strong selective reporting bias. The estimated constant represents the underlying mean partial correlation coefficient cleansed for the selective reporting bias. It is negative, but insignificant. It is supported by mean value of 0.00 reported in Table 3.1. Hence, the baseline result suggests that the evidence for the size effect is negligible in the data.

Table 5.1: Test of Publication Bias

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	0.844***	1.208**	0.813	1.028*	0.925***
	(0.155)	(0.552)	(0.623)	(0.563)	(0.181)
Constant	-0.012	-0.008	-0.011*	-0.002	-0.015
	(0.018)	(0.029)	(0.006)	(0.015)	(0.018)
N	522	522	522	522	522

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. In specification (1) OLS is used. Following specification (2) is panel data regression with between effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

The column (2) shows the result of panel data regression with between effects. The between effects indicate a selective reporting bias, which is even stronger than in case of OLS. The corrected partial correlation coefficient is again not significant. Moving on, I get to the analysis of WLS estimation in the column (3). Here the precision variable was used as weight. As shown in Table 5.1, only this specification pronounce publication bias insignificant and the mean partial correlation coefficient of -0.011 significant. The last two columns are linked to WLS estimation with the inverse number of the estimates reported per study as weight and to instrumental variable estimation. In the IV case, the inverse of the square root of the number of observations per study is used as an instrument. These two results correspond with conclusion of OLS and between effects. They report presence of selective reporting bias and negative but insignificant size effect.

Since above mentioned results of publication bias are for partial correlation coefficient, I decide to replicate all five approaches on unadjusted data for log-level and log-log case of return-volume relationship investigated in primary studies. The results for log-level case are summed up in Appendix A in Table A.3, and the conclusions for log-log data are summarized in the same place in Table A.4. In a nutshell, the results from both subsamples support the conclusions from the partial correlation coefficients about negative, but insignificant true mean and positive and significant selective reporting bias.

Moreover, I want to investigate whether publication bias is ceasing in newer studies or it is preserved phenomena. Besides I am curious if publication bias dominates among all journals, or only in the top ones. Therefore, I include interaction terms of the standard error with the year of publication and between standard error with the recursive impact factor in Equation 4.1. This approach employed Astakhov *et al.* (2017) or Havranek & Irsova (2012) among others. Results of only OLS specification from Table 5.1 are presented in the Table 5.2. Other specification brings the similar message and are presented in Appendix A in Table A.5 to Table A.7.

Table 5.2: Test of Publication Bias

	(1) OLS	(2) OLS	(3) OLS
SE	10.206**	1.842**	11.098***
	(3.235)	(0.793)	(3.633)
SE*Pubyear	-0.005**		-0.005***
	(0.002)		(0.002)
SE*Impact		-0.365	-0.363
		(0.270)	(0.270)
Constant	-0.011	-0.029	-0.028
	(0.018)	(0.020)	(0.019)
N	522	522	522

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \gamma * SE(S_{it}) * X_t + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. The X_t is either year of the publication of the study t, or an impact factor of the outlet, in which study t was published. The regressions are estimated by OLS with standard errors clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

The findings regarding the publication year are both significant at 5% level and negative, which suggest that the publication bias has been less of an issue in more recent studies. The direction is confirmed by the robustness check in Table A.5 and Table A.7, but it is considered insignificant. Similarly, the effect of journal quality on the publication bias is not confirmed, since both results are insignificant even at 10% level. Seemingly, journal quality does not influence the existence and extent of selective reporting bias.

Now I turn to results from the advanced techniques for publication bias detection and discuss them in brief. I start with results from Andrews & Maxmilian (2019) methodology. The corrected mean, in this case, is 0.013 with standard error 0.009. On the first glance, the corrected mean seems to have opposite sign than in case of "classical: approaches, on the other hand, the result is again insignificant, which confirms previous findings of true effect of negligible value. The same conclusion can be drawn from the adequate power technique, called WAAP, developed by Stanley & Doucouliagos (2012). Even after reduction of the requirements for adequately powered observation, there are just few such observations in the dataset, namely 36. Despite these mitigating facts, the result again support the findings of slightly negative but insignificant true mean. As the last technique for publication bias detection I mentioned the one suggested by Furukawa (2019). His stem-based estimation used 467 observations out of 522 in the contrast to only 36 employed by WAAP. Even this recent approach supports previous findings about negative-zero insignificant true effect.

5.2 Bayesian and Frequentist Model Averaging

Moving on to discussion of results of Bayesian Model Averaging and Frequentist Model Averaging I need to remind that since I cannot use the expected return-volume relationship directly, I use the partial correlation coefficient instead. Therefore, I am able only distinguish whether the control variables have positive or negative influence on the coefficient. The results of Bayesian Model Averaging and robustness check via Frequentist Model Averaging are both summed up in the following table - Table 5.3 in order to easier eye-check comparison of the results. I publish results of all variables and I divide them into several categories to better understanding of results[1].

The significance of variables in BMA approach is indicated as in Eicher *et al.* (2011). The division is based on posterior inclusion probability. The variable is considered as weak if the PIP is between 0.5 and 0.75. If the PIP is in range between 0.75 and 0.95, the variable is classified as substantial. With PIP between

[1]The data and source codes are available upon request.

0.95 - 0.99 the variable is named as strong. When the PIP is even above 0.99, the variable is categorized as decisive for the correct estimation.

I classified sixteen variables (without intercept) as decisive, which indicates that BMA is proper choice for estimation, since while one tries to use some the best model technique, some of the key variables could be easily omitted. Moreover, I observe high stability among the variables. It is easily visible from Figure 5.2. When the variables have constantly "blue boxes", it indicates a positive impact on the return-volume relationship across all models. When the variables have constantly "red boxes", it means that the variable is always negatively connected to the return-volume relationship. When there are changes between red and blue in relation to one specific variable, it means that the variable is unstable across the models.

The Figure 5.2 and Table 5.3 are attached and thoroughly discussed on the following pages.

Table 5.3: Results of Bayesian and Frequentist Model Averaging

	Bayesian Model Averaging			Frequentist Model Averaging		
Variable	*Pos. Mean*	*Pos. SD*	*PIP*	*Coef.*	*Std. Error*	*p-value*
Intercept	**0.001**	N/A	1.000	**0.003**	0.002	0.085
Se	**3.128**	0.344	1.000	**2.589**	0.384	0.000
Contem	**0.128**	0.011	1.000	**0.118**	0.012	0.000
Spillover	-0.031	0.088	0.154	-0.201	0.119	0.091
Return						
Absolute	0.006	0.020	0.127	0.022	0.036	0.545
Abnormal	0.013	0.034	0.193	0.030	0.047	0.525
Excess	-0.021	0.046	0.231	-0.068	0.061	0.269
Volume						
Volume	**0.036**	0.044	0.519	**0.071**	0.031	0.023
Shares	**-0.064**	0.030	0.878	**-0.028**	0.026	0.279
Detrend	0.001	0.006	0.062	0.008	0.029	0.775
Continent						
Asia	0.004	0.011	0.158	-0.012	0.029	0.685
Europe	-0.015	0.021	0.394	-0.073	0.033	0.025
Australia	0.000	0.008	0.057	-0.034	0.045	0.447

Continuation of Table 5.3						
	Bayesian Model Averaging			Frequentist Model Averaging		
Variable	*Pos. Mean*	*Pos. SD*	*PIP*	*Coef.*	*Std. Error*	*p-value*
Stock Exchange						
Index	-0.013	0.020	0.374	-0.047	0.028	0.090
Bank	**-0.274**	0.098	1.000	**-0.297**	0.097	0.002
Nonfin	0.003	0.016	0.080	0.000	0.052	0.998
Othinv	0.027	0.033	0.479	0.040	0.037	0.279
Type of Data						
Time	0.000	0.006	0.064	-0.087	0.037	0.017
Cross	0.032	0.043	0.424	0.006	0.061	0.926
Frequency						
Hourly	0.020	0.044	0.238	0.190	0.060	0.002
Daily	**0.115**	0.026	0.996	**0.155**	0.030	0.000
Weekly	-0.001	0.016	0.071	0.140	0.060	0.020
Dataset						
Length	**-0.005**	0.003	0.822	**-0.001**	0.002	0.752
Midyear	**-0.003**	0.002	0.818	**0.000**	0.002	0.950
Deving	**-0.077**	0.020	1.000	**-0.066**	0.024	0.005
Total	**0.043**	0.011	1.000	**0.029**	0.009	0.001
Msize	**-0.051**	0.006	1.000	**-0.058**	0.008	0.000
Model						
FamMac	**-0.444**	0.054	1.000	**-0.457**	0.078	0.000
Vara	0.083	0.099	0.483	0.125	0.092	0.174
Varb	-0.004	0.016	0.129	0.076	0.058	0.189
Varc	0.003	0.013	0.112	0.104	0.055	0.059
Simple	**0.110**	0.020	1.000	**0.144**	0.029	0.000
Garch	0.014	0.025	0.308	0.110	0.055	0.046
OthMod	-0.007	0.025	0.120	-0.071	0.051	0.163
Dummies						
Monday	-0.011	0.031	0.154	-0.214	0.059	0.000
Yld	**0.516**	0.123	0.989	**0.520**	0.144	0.000
Stdev	0.016	0.040	0.191	0.118	0.059	0.046
MarBet	-0.008	0.026	0.144	-0.064	0.042	0.125

Continuation of Table 5.3						
	Bayesian Model Averaging			Frequentist Model Averaging		
Variable	*Pos. Mean*	*Pos. SD*	*PIP*	*Coef.*	*Std. Error*	*p-value*
Illiq	-0.010	0.044	0.103	-0.280	0.116	0.016
Accrual	**0.376**	0.101	0.990	**0.534**	0.133	0.000
SaleP	-0.002	0.023	0.065	-0.249	0.099	0.012
FirmBet	0.011	0.057	0.083	0.206	0.149	0.167
Size	**-0.342**	0.059	0.997	**-0.129**	0.081	0.111
Trimmed	0.001	0.026	0.050	0.063	0.106	0.551
ExclJan	0.004	0.031	0.057	0.041	0.115	0.724
Informed	0.005	0.024	0.073	0.078	0.065	0.229
Ivol	**0.078**	0.044	0.847	**0.086**	0.053	0.103
Methodology						
Mle	-0.004	0.046	0.053	-0.020	0.159	0.901
Gmm	0.004	0.012	0.168	-0.048	0.045	0.286
Othest	-0.001	0.007	0.088	-0.074	0.041	0.072
Publishing						
Impact	**-0.148**	0.045	0.984	**-0.197**	0.053	0.000
Cit	0.000	0.002	0.084	0.008	0.006	0.220
Pblshd	0.001	0.006	0.058	0.036	0.035	0.305
Observations	522			522		

Notes: This table shows results from Bayesian Model Averaging and Frequentist Model Averaging as described in Section 4.2 and Section 4.3, respectively. The variables in bold are those with PIP (Posterior Inclusion Probability) from BMA larger than 0.5.

At first glance I see that the *Intercept* is classified as decisive variable in BMA with positive relation (+0.001) to expected return-volume partial correlation coefficient, which is of the main interest. Similarly, it is in case of FMA, when the positive sign (+0.003) remains and according to p-value, the inclusion of *Intercept* is significant at 10% level. This result, together with significant positive sign of *Standard Error* (PIP = 1.0, p-value = 0.0) support the conclusion of publication bias among the reported estimates in this research area as it is concluded in the discussion of publication bias results. Moreover, it supports the conclusion from Section 5.1 of negligible value of expected return-volume relationship. The same is proposed by guideline written by Doucouliagos (2011) as I mentioned in Subsection 3.1.1. This is in line with findings of Statman *et al.* (2006) from 2,000 US stocks, or Gurgul *et al.* (2007) on the aggregate level and Lee & Rui (2002), or Vo (2017) on the individual level. Furthermore, it can explain, why I do not observe any positive correlation between returns and volume in the futures markets (Chen *et al.* 2001).

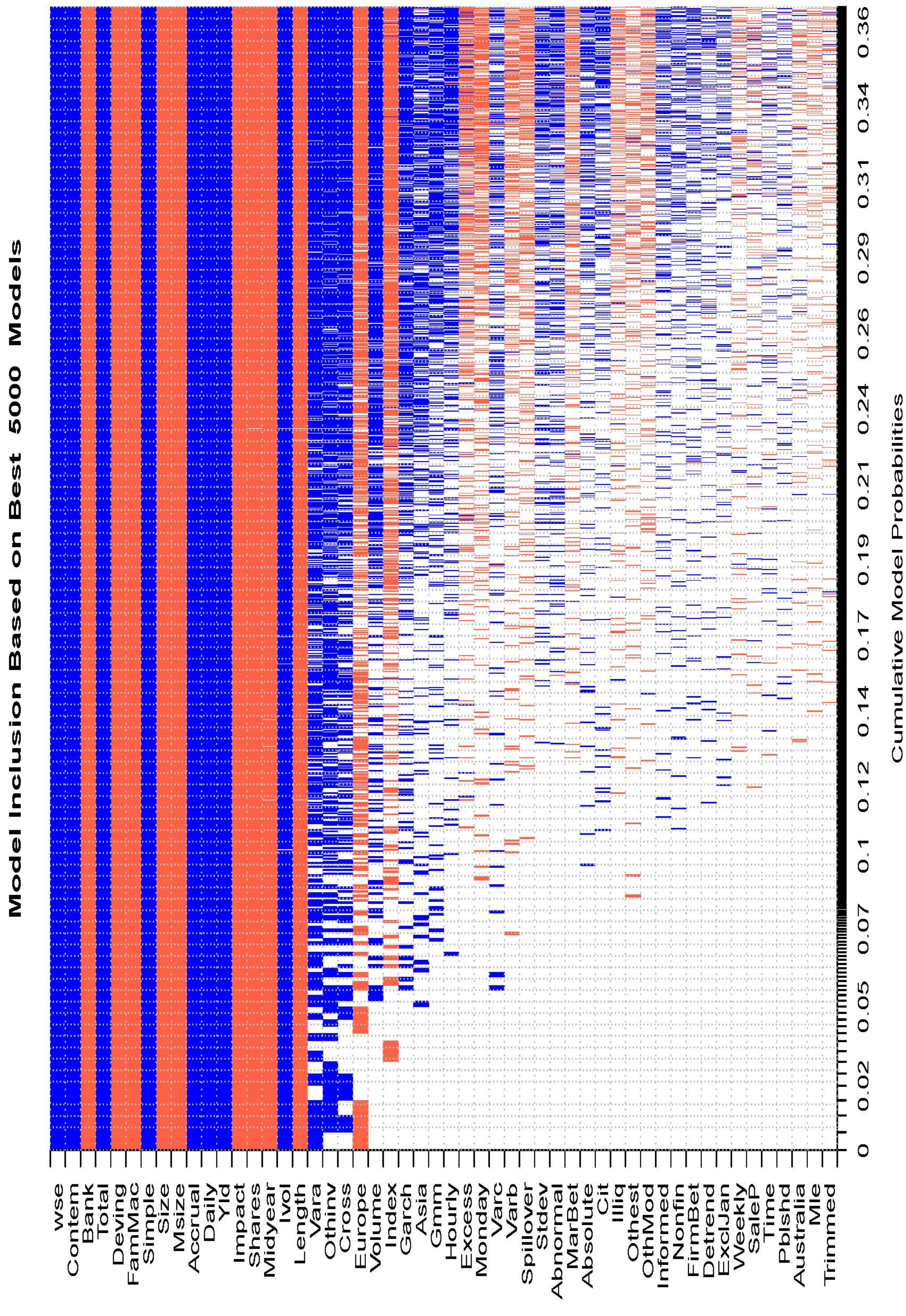

Figure 5.2: Bayesian Model Averaging

Besides I find that relationship between returns and volume is stronger in case, when the returns and volume are from the same time period. In this case the increase in volume about one unit causes increase in volume for 0.128 unit according to BMA model and about 0.118 in case of FMA robustness check. Both estimates are significant. Again, it is in line with the majority of primary studies. The same find, for example, Epps & Epps (1976), or Lee & Rui (2000).

On the contrary, the difference between intramarket return-volume connection and intermarket *Spillover* effect is not significant from Bayesian point of view. The FMA finds some negative effect (-0.201) in this case, which cannot be rejected at 10% level, but in case of 5% level it might be rejected easily, thus I will not dedicate much attention to it. The *Spillover* effect is investigated, for instance, by Kim (2005).

When I turn to discussion of distinctions among the different attitudes to return measure I do not find any significant differences among different approaches. It seems that the approach to return estimation is of negligible importance. The same is not true in case of different measures of trading volume authors who use dollar share volume as a measure of trading volume report substantial higher estimated coefficient (+0.036 in case of BMA and +0.071 in case of FMA) than authors who use turnover as a measure. On the other hand, it ought to be mentioned that the variable *Volume* is considered as weak, so the difference is captured just by half of the models. On the contrary, the results produced by data using raw number of shares as a volume measure shows even lower estimated coefficient than the data using share turnover. But this relationship is significant in BMA only with the posterior mean -0.064 and PIP 0.878, in FMA the hypothesis of no difference between employing share turnover and raw share volume cannot be rejected at 10% level. The last variable in differentiating the volume measures capture, if the primary volume data was *Detrend*ed or not. Whether the authors employed this procedure or not is found insignificant.

Moving on, I get to discussion of datasets used. At first glance, I can claim that the continent does not influence results. Just *Europe* in FMA model is significant at 5% level, but the BMA do not support such conclusion. When I consider, whether authors aggregate data and employed all shares from particular stock exchanges, or use just some index or use disaggregate data on individual level or in some particular sector I see no significant difference except the banking sector. Obviously the relationship between trading volume a expected stock returns is weak in the banking sector and it might be say it is negative. The estimated coefficient is -0.274 and the variable *Bank*, which captures the relationship is categorized as decisive. The same is supported by FMA with significance at 1% level. The result of no difference between aggregated and disaggregated data is something I expected based on findings from Lee & Rui (2000) and Lee & Rui (2002). On the other hand, the negative relationship in the *Banking* sector is something I do not recognize at the first glance and even after reverse check I do not find any counterweight in studies (Al-Jafari & Tliti 2013, Rotila *et al.* 2015) on *Banking* sector to this finding.

Similarly, the different type of data employed in primary data does not influence the estimated coefficient. The *Cross*-sectional, *Time* series and *Panel* data produce the similar results. Almost the same holds for different time frequency. The significantly different results are found in case of using *Daily* frequency of data gathering. The estimates using *Daily* frequency show substantially higher results (+0.115 in BMA and +0.155 in FMA) than the estimates using *Monthly* frequency. This might explain why the results of *Time* series data and *Detrend*ing are insignificant. The *Daily* frequency is usually connected with *Time* series data, which are *Detrend*ed. This is not a rule, but it is common practice, used, for instance, by Yonis (2014), or Le & Mehmed (2009). So the difference is captured by BMA and FMA in data frequency and the type of data and *Detrend*ing are found to be insignificant. In FMA is this idea supported even by frequency variable *Hourly*, which has positive effect (+0.190) and it is significant at 1% level.

The largest differences in primary datasets are captured by variables *Length*, *Midyear*, *Deving*, *Total* and *Msize*. The impact of variables *Length* (-0.005) and *Midyear* (-0.001) seems negligible, but I need to realize that the *Length*, which is computed in year varies from few months (Epps & Epps 1976) to 52,5 years (Han *et al.* 2018), so the final difference can vary up to -0.263. Similarly, the difference of the oldest dataset according to *Midyear* (Crouch 1970) and the newest (Tapa & Hussein 2016) is about 45 years, which could cause difference up to -0.135. This again notable. On the other hand, FMA does not support any of these conclusions. Both estimation approaches support the hypothesis about different relationship between trading volume and expected stock returns in developing and developed countries. The estimated coefficients are substantially smaller for the developing countries (variable *Deving*). Next differences are caused by variables in logarithms - *Total* and *Msize*. While variable *Total* suggests higher estimated coefficient between expected stock returns and trading volume with increasing number of observations, the variable *Msize* suggests lower estimations connected to larger markets. These variables might explain insignificant difference among variables *Cross*-sectional, *Time* series and *Panel*, since the *Panel* data are those with the highest number of observations. Similarly, the variable *Msize* helps me to understand of no meaningful differences among continents, since the largest market according to GDP is US market (especially in recent years) and the smallest are those from developing countries. This findings counterweight the results of *Deving* (country is either developing or has larger GDP, both has negative effect) and amplify the effect of *Midyear* (newer datasets have larger market size value, both with negative effect).

Now I move to different models used in primary studies. While the Fama-Macbeth (*FamMac*) specification provides substantially lower results, the model category *Simple* provides substantially higher results than the default specification. It is not surprising, since Fama-Macbeth usually employ panel data and the *Simple* specification is particularly used with time series. Thus these two variables stands against variable *Total* number of observation, which draw the result in the opposite direction. Other differences are caused by so called *Dummies*. This variables capture anomalies employed by low number of studies, but above threshold (at least 2% of observation). So this result affects some individual studies, not whole population. Among significant variables with positive sign, I rank the yield (*Yld*), *Accrual* and *Size*. The value of *Yld* seems large in comparison with effects of other variables, but I have to realize that I recognize this variable only in two studies (Brennan *et al.* 1998 and Chordia *et al.* 2001) and these two researches employs *FamMac* model, which almost counterweight the effect of the yield (0.516 and -0.444). Similarly, it is in case of variable *Size*, which stands against the model used defined as *Simple* (-0.342 and 0.110) and supports the effect of *Total* number of observations as mentioned in the previous paragraph. On the contrary, I find interesting the variable *Accrual*, which shows whether or not the primary article considers the firm's change in non-cash net working capital minus depreciation in the prior fiscal year. This variable is employed only by Lin & Liu (2017) and Lewellen (2015). After further investigation, I conclude that this variable might be balance out by the *Length* of the data set, since both articles use long data; 39 and 46 years, respectively. What is true about different models employed by authors, does not hold in case of estimation methodology. Neither BMA, nor FMA find any significant explanatory variable in this area.

The last notes of results discussion are dedicated to the published and unpublished articles. The explanatory variable *Pblshd*, which differentiate between the published and unpublished articles and the variable *Cit*ations, which number the count of citations of the articles are of no importance. The same is not true about the variable *Impact*. It seems like the results published in more influential journals are substantially lower on average. This assertion is backed up by both methodologies BMA (-0.148, PIP = 0.984) and FMA (-0.197, significant at 1% level).

Chapter 6

Conclusions

I conduct a meta-analysis of the relationship between expected stock return and trading volume. I collect 522 estimates from 46 studies. I control for numerous differences, finally for 52, such as midyear of data, type of data, length of the primary dataset, market size or model used in primary study.

Besides all interesting findings related to the expected stock returns - trading volume relationship itself, I investigate the publication selection bias in this area. To enhance the reliability of the results, I do not employ only common approaches as funnel plot or formal tests using OLS, between effects, WLS or instrumental variables estimation, but I employ several recent approaches. These new approaches such as Andrews & Maxmilian (2019) estimator, WAAP (Stanley *et al.* 2017) or stem-based estimator (Furukawa 2019) were introduced recently since the public often questioned the results from common formal tests. During my thorough investigation of publication bias regardless of the technique employed I come to two convincing results. The size effect corrected for the publication bias is of negligible value and is insignificant. Furthermore, this is later confirmed by results from BMA and FMA itself. The second finding from these researches show that there is evidence across the techniques that the publication selection bias is present in the data. This is again validated in BMA and its robustness check. Moreover, I find out that it is not effected by the quality of journal, which reports the studies and that the hypothesis of publication bias decreasing in time is again not so clear in this field.

In investigation of the expected stock returns - trading volume relationship itself I employ commonly used Bayesian model averaging (BMA) approach, but in addition to BMA, I use modern Frequentist model averaging (FMA). BMA provides weighted average of estimates of explanatory variables based on large number of models using Metropolis-Hastings algorithm. On the contrary, FMA approach uses the weighted average of the best models as well, but this selection of the models is based on the goodness of fit and parsimony of potentially included models. To make the model averaging feasible, FMA employs orthogonalization of the variable space.

Finally, I find that there is a small or even negligible relationship between trading volume and expected stock returns as it is proposed by many primary studies such as Mahajan & Singh (2009a) and Akpansung & Gidigbi (2015). On the other hand, my result contradict findings of others (Hafner 2005, Hu 1997). In contemporaneous terms, the relationship between trading volume and stock returns is more obvious. The contemporaneous effect augments the estimated coefficient by 0.128. Again, it is something what several primary studies suggest, for instance, Epps & Epps (1976), or Lee & Rui (2000), but again others find the opposite (Chordia *et al.* 2001, Sheu *et al.* 1998). Therefore, my results may serve as a baseline in the model calibration, or can be used directly in traders' strategies.

Besides the investigation of the heterogeneity among primary studies brings several other interesting notes. For instance, in bank sector, which is investigated by Al-Jafari & Tliti (2013) and Rotila *et al.* (2015), there is negative impact on the relationship between trading volume and expected stock returns. On the other side, when the authors collect the data on the hourly or daily basis, the relationship here is much more convincing than between different months periods. In addition to this, other relationships are found, for example, there is no difference in results caused by level of aggregation used by different authors. The same holds for type of data, all three types; time series, cross-sectional and panel data produce comparable results. Another interesting thing is that estimated coefficients from developing and smaller countries are substantially lower then those for developed and large countries. The estimated effects are -0.077 and -0.051, respectively, and both are statistically significant. Other interesting finding comes from the investigation among the studies published in impacted journals. It seems like the results published in the more influential journals are lower on average, than other estimates. The direction of this finding is in line with the one from the publication bias section, but in that case it was insignificant. Therefore, I should be cautious to do any impetuous conclusions.

Bibliography

Akpansung, A. O. & M. O. Gidigbi (2015): The Relationship between Trading Volumes and Returns in the Nigerian Stock Market. *International Research Journal of Finance and Economics* 132: pp. 150–163.

Al-Jafari, M. K. & A. Tliti (2013): An empirical investigation of the relationship between stock return and trading volume: Evidence from the Jordanian banking sector. *Journal of Applied Finance and Banking* 3(3): pp. 45–64.

Amini, S. M. & C. F. Parmeter (2012): Comparison of model averaging techniques: Assessing growth determinants. *Journal of Applied Econometrics* 27(5): pp. 870–876.

Andrews, I. & K. Maxmilian (2019): Identification of and Correction for Publication Bias. *American Economic Review* p. forthocoming.

Assogbavi, T., J. Schell, & S. Fagnisse (2007): Equity price-volume relationship on the Russian Stock Exchange. *International Business and Economics Research Journal* 6(9): pp. 107–116.

Astakhov, A., T. Havranek, & J. Novak (2017): Firm Size and Stock Returns: A Meta-Analysis. *Technical report*, Charles University.

Bajzik, J. (2017): A Meta-Analysis of the Estimates of the Armington Elasticity. *Technical report*, Charles University.

Brandle, A. (2010): Volume Based Portfolio Strategies - Analysis of the Relationship between Trading activity and Expected Returns in the Cross-Section of Swiss Stocks. *Technical report*, University of St. Gallen.

Brennan, M. J., T. Chordia, & A. Subrahmanyam (1998): Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Returns. *Journal of Financial Economics* 49: pp. 345–373.

Chang, S. S. & F. A. Wang (2019): Informed contrarian trades and stock returns. *Journal of Financial Markets* 42: pp. 75–93.

Chen, G.-m., M. Firth, & O. M. Rui (2001): The dynamic relation between stock returns, trading volume, and volatility. *Financial Review* 36(3): pp. 153–174.

Chordia, T., A. Subrahmanyam, & V. R. Anshuman (2001): Trading activity and expected stock returns. *Journal of financial Economics* 59(1): pp. 3–32.

Chuang, W.-I. & B.-S. Lee (2006): An empirical evaluation of the overconfidence hypothesis. *Journal of Banking and Finance* 30(9): pp. 2489–2515.

Ciccone, A. & M. Jarocinski (2010): Determinants of economic growth: Will data tell? *American Economic Journal: Macroeconomics* 2(4): pp. 222–246.

Ciner, C. (2002): The stock price-volume linkage on the Toronto stock exchange: Before and after automation. *Review of Quantitative Finance and Accounting* 19(4): pp. 335–349.

Ciner, C. (2003): Dynamic linkages between trading volume and price movements: Evidence for small firm stocks. *The Journal of Entrepreneurial Finance* 8(1): pp. 87–102.

Clark, P. K. (1973): A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society* pp. 135–155.

Cohen, J. (1988): Statistical power analysis for the behaviors science. *Working paper*, New Jersey: Laurence Erlbaum Associates, Publishers, Hillsdale.

Comiskey, E. E., R. A. Walkling, & M. A. Weeks (1987): Dispersion of Expectations and Trading Volume. *Journal of Business Finance and Accounting* pp. 229–239.

Copeland, T. (1976): A Model of Asset Trading under the Assumption of Sequential Information Arrival. *The Journal of Finance* 31: pp. 1149–1168.

Cornell, B. (1981): The relationship between volume and price variability in futures markets. *Journal of Futures Markets* 1(3): pp. 303–316.

Crouch, R. L. (1970): The volume of transactions and price changes on the New York Stock Exchange. *Financial Analysts Journal* 26(4): pp. 104–109.

Darrat, A. F., S. Rahman, & M. Zhong (2003): Intraday trading volume and return volatility of the DJIA stocks: A note. *Journal of Banking and Finance* 27(10): pp. 2035–2043.

Datar, V. T., N. Y. Naik, & R. Radcliffe (1998): Liquidity and stock returns: An alternative test. *Journal of Financial Markets* 1(2): pp. 203–219.

De Bondt, W. F. M. & R. Thaler (1985): Does the Stock Market Overreact? *The Journal of Finance* 40: pp. 793–805.

De Meiros, O. R. & B. F. N. Van Doornik (2008): The Empirical Relationship between Stock Returns, Return Volatility and Trading Volume in the Brazilian Stock Market. *Brazilian Business Review (English Edition)* 5(1).

Devanadhen, K., K. Srnivasan, & M. Deo (2010): Price Changes, Trading Volume and Time-Varying Conditional Volatility: Evidence from Asia Pacific Stock Market. *International Review of Applied Financial Issues and Economics* 2(2): p. 379.

Doucouliagos, C. (2005): Publication bias in the economic freedom and economic growth literature. *Journal of Economic Surveys* 19(3): pp. 367–387.

Doucouliagos, C. (2011): How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics. *Economics series working paper 05*, Deakin University.

Doucouliagos, H. & T. D. Stanley (2013): Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity. *Journal of Economic Surveys* 27(2): p. 316–339.

Durlauf, S. N., A. Kourtellos, & C. M. Tan (2008): Are Any Growth Theories Robust? *The Economic Journal* 118(527): pp. 329–346.

Egger, M., G. D. Smith, & C. Minder (1997): Bias in Meta-Analysis Detected by a Simple, Graphical Test. *Journal of Economic Surveys* 316: p. 629–634.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants. *Journal of Applied Econometrics* 26(1): pp. 30–55.

Epps, T. W. & M. L. Epps (1976): The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica: Journal of the Econometric Society* pp. 305–321.

Feldkircher, M., R. Horvath, & R. Marek (2014): Exchange market pressures during the financial crisis: A Bayesian model averaging evidence. *ournal of International Money and Finance* 40: pp. 21–41.

Feldkircher, M. & S. Zeugner (2009): Benchmark priors revisited: On adaptive shrinkage and the supermodel effect in bayesian model averaging. *Working Paper, International Monetary Fund* 09(202).

Fernandez, C., E. Ley, & M. F. J. Steel (2001): Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* 16(5): pp. 563–576.

Fisher, R. (1954): *Statistical Methods for Research Workers, 12th edn.* Edinburgh, UK: Oliver and Boyd.

Furukawa, C. (2019): Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method. *Technical report*, ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg.

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

Godfrey, M. D., C. W. Granger, & O. Morgenstern (1964): The random-walk hypothesis of stock market behavior. *Kyklos* 17(1): pp. 1–30.

Grammatikos, T. & A. Saunders (1986): Futures price variability: A test of maturity and volume effects. *Journal of Business* pp. 319–330.

Granger, C. W. (1969): Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* pp. 424–438.

Granger, C. W. & O. Morgenstern (1963): Spectral analysis of New York stock market prices 1. *Kyklos* 16(1): pp. 1–27.

Greene, W. H. (2008): *Econometric Analysis.* Upper Saddle River, New Jersey: Pearson International in Prentice-Hall International Editions.

Gurgul, H., P. Majdosz, & R. Mestel (2007): Price Volume Relations of DAX Companies. *Financial Markets and Portfolio Management* 21: pp. 353–379.

Hafner, C. M. (2005): Durations, volume and the prediction of financial returns in transaction time. *Quantitative Finance* 5(2): pp. 145–152.

Han, Y., D. Huang, D. Huang, & G. Zhou (2018): Volume and Return: The Role of Mispricing. *Ssrn*, University of North Carolina.

Hansen, B. E. (2007): Least squares model averaging. *Econometrica* 75(4): p. 1175–1189.

Harris, L. (1987): Transaction Data Tests of the Mixture of Distributions Hypothesis. *Journal of Financial and Quantitative Analysis* 22: pp. 127–141.

Haugen, R. A. & N. L. Baker (1996): Commonality in the Determinants of Expected Stock Returns. *Journal of Financial Economics* 41: pp. 401–439.

Havranek, T. & Z. Irsova (2012): Survey Article: Publication Bias in the Literature on Foreign Direct Investment Spillovers. *Journal of Development Studies* 48(10): pp. 1375–1396.

Havranek, T. & Z. Irsova (2015): Do borders really slash trade? A meta-analysis. *IMF Economic Review* pp. 1–32.

Havranek, T., M. Rusnak, & A. Sokolova (2017): Habit formation in consumption: A meta-analysis. *European Economic Review, Elsevier* 95(C): pp. 142–167.

Havranek, T. & A. Sokolova (2019): Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say 'Probably Not'. *Working paper*, Czech National Bank and Charles University, Prague.

Horvath, R., E. Horvatova, & M. Siranova (2017): Financial development, rule of law and wealth inequality: Bayesian model averaging evidence. *IOS Working Papers* 368.

Hu, S.-Y. (1997): Trading Turnover and Expected Stock Returns: The Trading Frequency Hypothesis and Evidence from the Tokyo Stock Exchange. *Working paper*, University of Chicago.

Jain, P. C. & G.-H. Joh (1988): The Dependence between Hourly Prices and Trading Volume. *The Journal of Financial and Quantitative Analysis* 23: pp. 269–283.

Jennings, R. H., L. T. Starks, & F. J. C. (1981): An equilibrium model of asset trading with sequential information arrival. *The Journal of Finance* 36(1): pp. 143–161.

Karpoff, J. M. (1987): The relation between price changes and trading volume: A survey. Journal of Financial and quantitative Analysis. *Journal of Financial and quantitative Analysis* 22(1): pp. 109–126.

Kim, S.-J. (2005): Information leadership in the advanced Asia–Pacific stock markets: Return, volatility and volume information spillovers from the US and Japan. *Journal of the Japanese and International Economies* 19(3): pp. 338–365.

KOCAGIL, A. E. & Y. SHACHMUROVE (1998): Return-Volume Dynamics in Futures Markets. *The Journal of Futures Market* 18(4): pp. 399–426.

LE, Q. T. & M. MEHMED (2009): The relationship between trading volume, stock index returns and volatility: Empirical evidence in Nordic countries. *Master thesis in finance*, Lund University.

LEE, C. & O. RUI (2000): Does Trading Volume Contain Information to Predict Stock Returns? Evidence from China's Stock Markets. *Review of Quantitative Finance and Accounting* 14: pp. 341–360.

LEE, C. & O. RUI (2002): "The Dynamic Relationship between Stock Returns and Trading Volume: Domestic and Cross-Country Evidence. *Journal of Banking and Finance* 26: pp. 51–78.

LEE, C. M. C. & B. SWAMINATHAN (2000): Price Momentum and Trading Volume. *The Journal of Finance* 55: pp. 2017–2069.

LEWELLEN, J. (2015): The cross section of expected stock returns. *Critical Financial Review* 4(1): pp. 1–44.

LIANG, F., R. PAULO, G. MOLINA, & B. J. O. CLYDE, Merlise A. (2008): Mixtures of g priors for bayesian variable selection. *Journal of the American Statistical Association* 103(481): pp. 410–423.

LIN, T.-C. & X. LIU (2017): Skewness, individual investor preference, and the cross-section of stock returns. *Review of Finance* pp. 1–36.

LO, A. W. & J. WANG (2000): Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory. *The Review of Financial Studies* 13: pp. 257–300.

LONG, H., Y. JINAG, & Y. ZHU (2018): Idiosyncratic tail risk and expected stock returns: Evidence from the Chinese stock markets. *Finance Research Letters* 24: pp. 129–136.

LOUHICHI, W. (2012): Does trading activity contain information to predict stock returns? Evidence from Euronext Paris. *Applied Financial Economics* 22(8): pp. 625–632.

LOUKIL, N., M. B. ZAYANI, & A. OMRI (2010): Impact of liquidity on stock returns: an empirical investigation of the Tunisian stock market. *Macroeconomics and Finance in Emerging Market Economies* 3(2): pp. 261–283.

Magnus, J. R., O. Powell, & P. Prufer (2010): A comparison of two model averaging techniques with an application to growth empirics. *Journal of econometrics* 154(2): pp. 139–153.

Mahajan, S. & B. Singh (2008): An empirical analysis of stock price-volume relationship in Indian stock market. *Vision* 12(3): pp. 1–13.

Mahajan, S. & B. Singh (2009a): The empirical investigation of relationship between return, volume and volatility dynamics in Indian stock market. *Eurasian Journal of Business and Economics* 2(4): pp. 113–137.

Mahajan, S. & B. Singh (2009b): Return-Volume Dynamics in Indian Stock Market. *Asia-Pacific Business Review* 5(3): pp. 63–70.

Marshall, B. R. & M. Young (2003): Liquidity and stock returns in pure order-driven markets: evidence from the Australian stock market. *International Review of Financial Analysis* 12(2): pp. 173–188.

McGowan, C. B. & J. Muhammad (2012): The relationship between price and volume for the Russian trading system. *The International Business and Economics Research Journal (Online)* 11(9): pp. 963–970.

Mcmillan, D. & A. Speight (2002): Return-Volume Dynamics in UK Futures. *Applied Financial Economics* 12: pp. 707–713.

Mestel, R., H. Gurgul, & P. Majdosz (2003): The empirical relationship between stock returns, return volatility and trading volume on the Austrian stock market. *Research paper*, University of Graz, Institute of Banking and Finance.

Moral-Benito, E. (2011): Model averaging in economics. *Banco de Espana Working Papers* 1123.

Narayan, P. K. & X. Zheng (2010): Market liquidity risk factor and financial market anomalies: Evidence from the Chinese stock market. *Pacific-Basin finance journal* 18(5): pp. 509–520.

Ochere, O. G., M. T. Nasieku, & T. O. Olweny (2018): Trading Volume and Fama-French Three Factor Model on Excess Return. Empirical Evidence from Nairobi Security Exchange. *European Scientific Journal* 14(22): pp. 276–289.

Pisedtasalasai, A. & A. Gunasekarage (2007): Causal and dynamic relationships among stock returns, return volatility and trading volume: Evidence from emerging markets in South-East Asia. *Asia-Pacific Financial Markets* 14(4): pp. 277–297.

Raftery, A. E. (1995): Bayesian model selection in social research. *Sociological methodology* pp. 111–163.

Raftery, A. E., D. Madigan, & J. A. Hoeting (1997): Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association* 92(437): pp. 179–191.

Richardson, G., S. E. Sefcik, & R. Thompson (1986): A test of dividend irrelevance using volume reactions to a change in dividend policy. *Journal of Financial Economics* 17(2): pp. 313–333.

Rotila, D.-M., M. Onofrei, & A. M. Andries (2015): The relation between stock returns, trading volume and return volatility of the CEE banks. *Transformations in Business and Economics* 14.

Rusnak, M., T. Havranek, & R. Horvath (2013): How to solve the price puzzle? A meta-analysis. *Journal of Money, Credit and Banking* 45(1): p. 37–70.

Saatcioglu, K. & L. T. Starks (1998): The stock price–volume relationship in emerging stock markets: the case of Latin America. *International Journal of forecasting* 14(2): pp. 215–225.

Sana Hsieh, H.-C. (2014): The causal relationships between stock returns, trading volume, and volatility: Empirical evidence from Asian listed real estate companies. *International Journal of Managerial Finance* 10(2): pp. 218–240.

Schürenberg-Frosch, H. (2015): We Could Not Care Less About Armington Elasticities–But Should We? A Meta-Sensitivity Analysis of the Influence of Armington Elasticity Misspecification on Simulation Results. *Working paper 594*, Ruhr Economic Papers.

Sheu, H.-J., S. Wu, & K.-P. Ku (1998): Cross-sectional relationships between stock returns and market beta, trading volume, and sales-to-price in Taiwan. *International Review of Financial Analysis* 7(1): pp. 1–18.

Shu, P.-G., Y.-H. Yeh, & Y.-C. Huang (2004): Stock price and trading volume effects associated with changes in the MSCI free indices: evidence from Taiwanese firms added to and deleted from the indices. *Review of Pacific Basin Financial Markets and Policies* 7(04): pp. 471–491.

Siegfried, J. J. (2012): Minutes of the Meeting of the Executive Committee: Chicago, IL, January 5, 2012. *American Economic Review* 102(3): pp. 645–52.

Stanley, T. D. & H. Doucouliagos (2012): *Meta-regression analysis in economics and business.* New York, USA: Routledge.

Stanley, T. D. & H. Doucouliagos (2014): Meta-Regression Approximations to Reduce Publication Selection Bias. *Research Synthesis Methods* 5(1): p. 60–78.

Stanley, T. D., H. Doucouliagos, & J. P. Ioannidis (2017): Finding the power to reduce publication bias. *Statistics in medicine* 36(10): pp. 1580–1598.

Statman, M., S. Thorley, & K. Vorkink (2006): Investor overconfidence and trading volume. *The Review of Financial Studies* 19(4): pp. 1531–1565.

Tahir, S. H., F. All, N. Ghaffar, & H. M. Sabir (2016): Causal relationship among trading volume, returns and stock volatility: evidence from an emerging market. *Technical report*, University Faisalabad.

Tapa, A. & M. Hussein (2016): The Relationship between Stock Return and Trading Volume in Malaysian ACE Market. *International Journal of Economics and Financial Issues* 6(75): pp. 271–278.

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

Tripathy, N. (2011): The relation between price changes and trading volume: A study in Indian stock market. *Interdisciplinary Journal of Research in Business* 1(7): pp. 81–95.

Valickova, P., T. Havranek, & R. Horvath (2015): Financial development and economic growth: A meta-analysis. *Journal of Economic Surveys* 29(3): pp. 506–526.

Vo, X. V. (2017): Trading of foreign investors and stock returns in an emerging market-Evidence from Vietnam. *International Review of Financial Analysis* 52: pp. 88–93.

Watkins, B. D. (2007): The Economic and Predictive Value of Trading Volume Growth: A Tale of Three Moments. *Applied Financial Economics* 17: pp. 1489–1509.

Wood, R. A., T. H. McInish, & J. K. Ord (1985): An Investigation of Transactions Data for NYSE Stocks. *The Journal of Finance* 60: pp. 723–739.

Wooldridge, J. M. (2015): *Introductory econometrics: A modern approach.* Scarborough, Canada: Nelson Education.

Yin, Y. & Y. Liu (2018): Information of Unusual Trading Volume. *Emerging Markets Finance and Trade* 54(11): pp. 2409–2432.

Ying, C. C. (1966): Market Prices and Volumes of Sales. *Econometrica* 34: pp. 676–685.

Yonis, M. (2014): Trading Volume and Stock Return: Empirical Evidence for Asian Tiger Economies. *Technical report*, UMEA Universitet.

Zeugner, S. (2011): Bayesian Model Averaging with BMS for BMS version 0.3. 0. *Online: www.bms.zeugner.eu*.

Zhong, A., D. Chai, B. Li, & M. Chiah (2018): Volume shocks and stock returns: An alternative test. *Pacific-Basin Finance Journal* 48: pp. 1–16.

Zigraiova, D. & T. Havranek (2016): Bank Competition And Financial Stability: Much Ado About Nothing? *Journal of Economic Surveys* 30(5): pp. 944–981.

Appendix A

Data

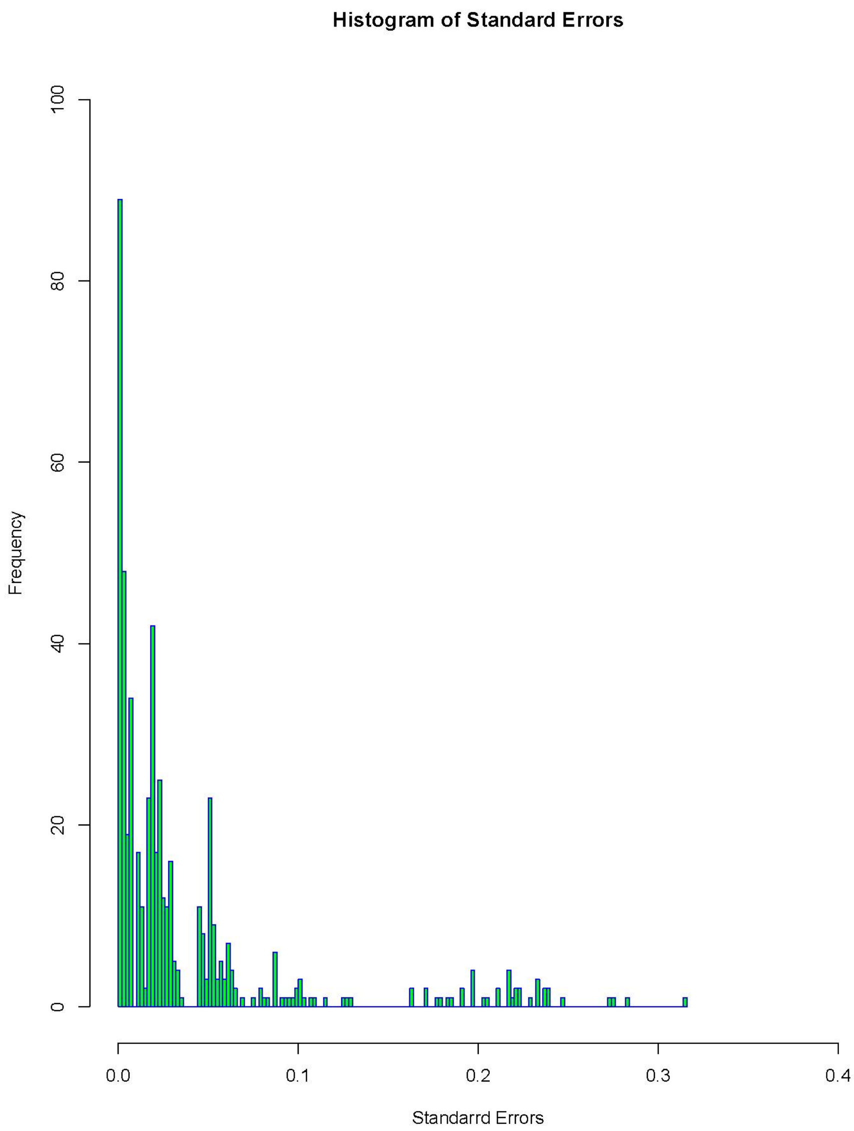

Figure A.1: Histogram of Standard Errors

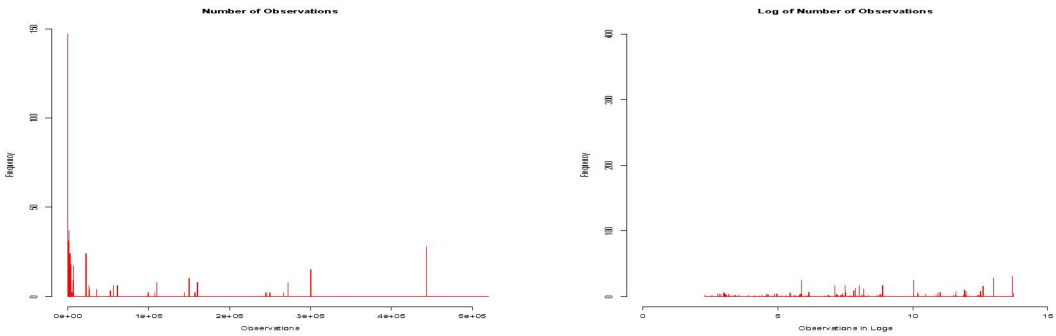

Figure A.2: Difference between distribution of total number of observations and their values in logs

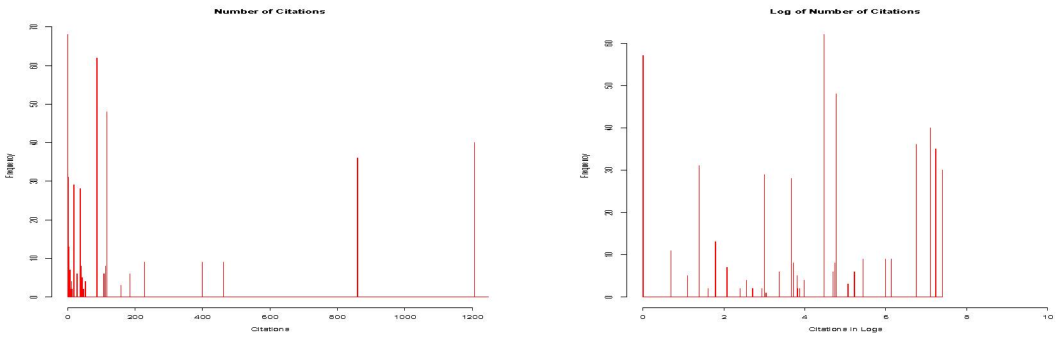

Figure A.3: Difference between distribution of number of citations and their values in logs

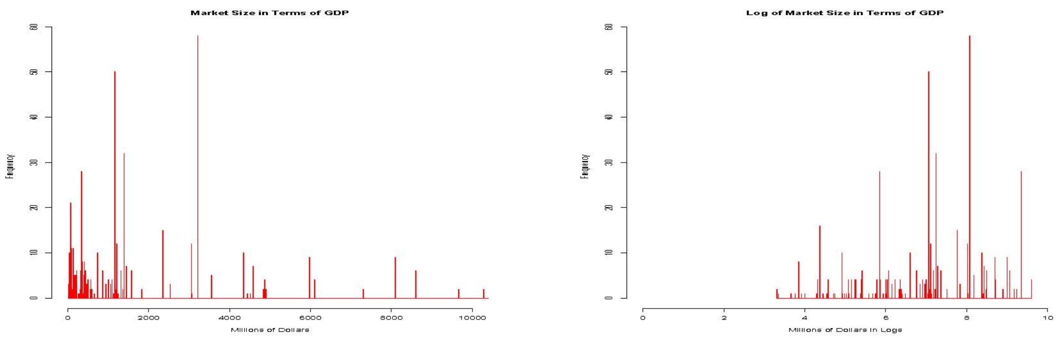

Figure A.4: Difference between distribution of Market sizes and their values in logs

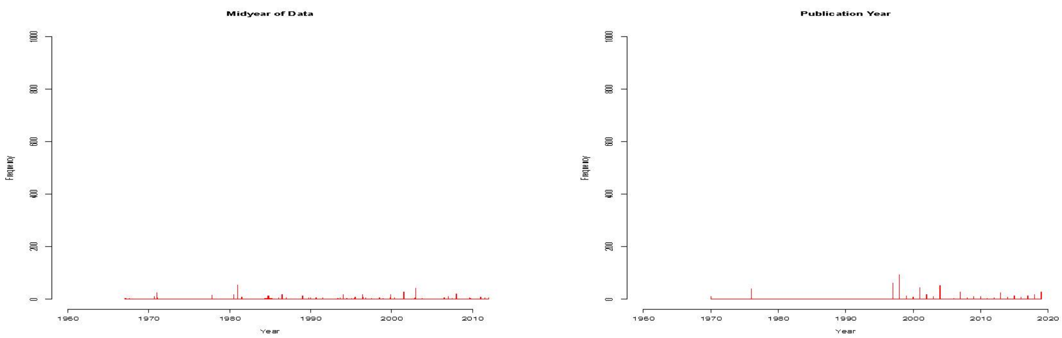

Figure A.5: Distribution of midyears of data and publication years

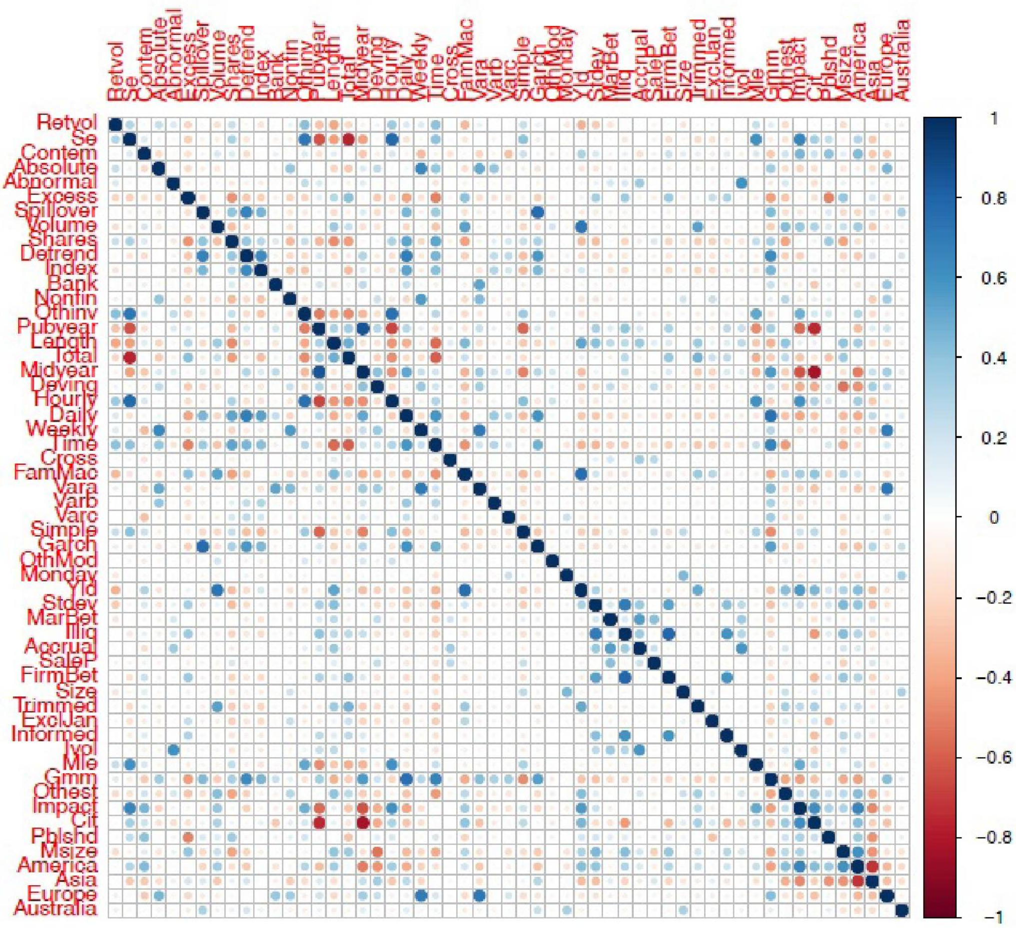

Figure A.6: Correlation Matrix

Table A.1: Correlation Matrix

Variable 1	*Variable 2*	*Correlation*
Pubyear	Midyear	0.854
Midyear	Cit	-0.823
Illiq	FirmBet	0.786
Se	Hourly	0.773
Se	Total	-0.767
Spillover	Garch	0.765

Notes: Only six correlation coefficients with the highest absolute value are presented. Based on this results I decide to withdraw the *Pubyear*.

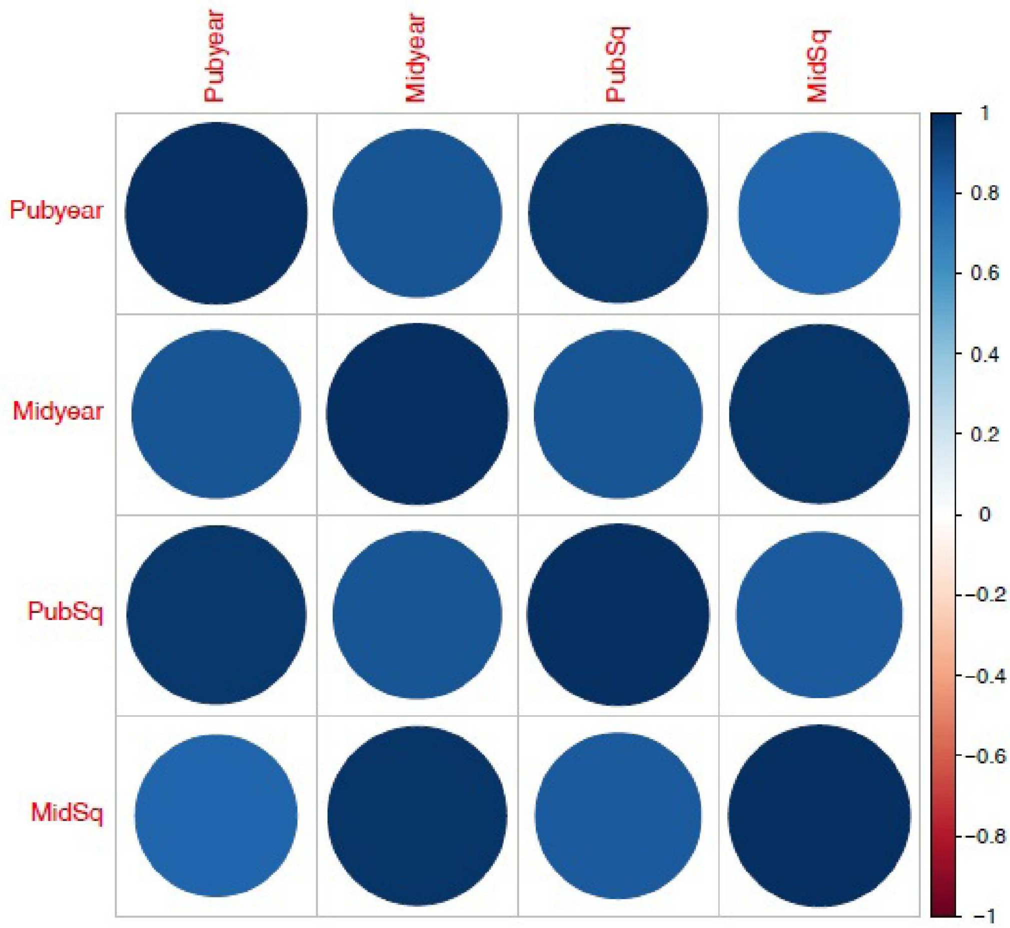

Figure A.7: Correlation Matrix between Publication Year, Midyear of Data and its Square Terms

Table A.2: Correlation Matrix

Variable 1	*Variable 2*	*Correlation*
Midyear	MidSq	0.974
Pubyear	PubSq	0.967
Midyear	PubSq	0.854
Pubyear	Midyear	0.854
PubSq	MidSq	0.836
Pubyear	MidSq	0.793

Notes: All six correlation coefficients are presented. Based on this results I decide to withdraw the square terms.

Table A.3: Test of Publication Bias for log-level cases

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	1.431***	0.591*	1.151***	1.000***	1.187
	(0.194)	(0.302)	(0.241)	(0.351)	(1.164)
Constant	-0.083	0.162	-0.000	0.048	-0.011
	(0.125)	(0.183)	(0.000)	(0.113)	(0.350)
N	279	279	279	279	279

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. In specification (1) OLS is used. Following specification (2) is panel data regression with random effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A.4: Test of Publication Bias for log-log cases

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	2.408**	3.469***	0.919	2.640**	3.634***
	(1.215)	(0.475)	(1.191)	(1.202)	(1.178)
Constant	-0.148	-0.244	-0.003**	-0.093	-0.268
	(0.099)	(0.208)	(0.001)	(0.061)	(0.205)
N	223	223	223	223	223

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. In specification (1) OLS is used. Following specification (2) is panel data regression with random effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A.5: Test of Publication Bias with Publication Year

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	10.206**	105.424	-17.458	17.059	308.248
	(3.235)	(68.729)	(41.378)	(14.649)	(198.160)
SE*Pubyear	-0.005**	-0.052	0.009	-0.008	-0.154
	(0.002)	(0.034)	(0.021)	(0.007)	(0.099)
Constant	-0.011	0.008	-0.011*	0.000	0.006
	(0.018)	(0.031)	(0.006)	(0.015)	(0.025)
N	522	522	522	522	522

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \gamma * SE(S_{it}) * P_t + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. The P_t is year of the publication of the study t. In specification (1) OLS is used. Following specification (2) is panel data regression with between effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A.6: Test of Publication Bias with Impact Factor

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	1.842**	1.706**	1.498**	1.483	2.35**
	(0.793)	(0.757)	(0.757)	(1.052)	(0.934)
SE*Impact	-0.365	-0.313	-0.414	-0.289	-0.522*
	(0.270)	(0.325)	(0.356)	(0.359)	(0.294)
Constant	-0.029	-0.019	-0.013**	-0.012	-0.039*
	(0.020)	(0.032)	(0.007)	(0.021)	(0.022)
N	522	522	522	522	522

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \gamma * SE(S_{it}) * I_t + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. The I_t is an impact factor of the outlet, in which study t was published. In specification (1) OLS is used. Following specification (2) is panel data regression with between effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.

Table A.7: Test of Publication Bias with Publication Year and Impact Factor

	(1) OLS	(2) BE	(3) Precision	(4) Study	(5) IV
SE	11.098***	214.336***	-12.576	20.149	305.604*
	(3.633)	(79.951)	(39.520)	(14.989)	(159.989)
SE*Pubyear	-0.005***	-0.106***	0.007	-0.009	-0.152*
	(0.002)	(0.040)	(0.020)	(0.008)	(0.080)
SE*Impact	-0.363	-0.877**	-0.384	-0.320	-0.858**
	(0.270)	(0.371)	(0.341)	(0.357)	(0.432)
Constant	-0.028	-0.007	-0.013**	-0.011	-0.034
	(0.019)	(0.030)	(0.007)	(0.021)	(0.026)
N	522	522	522	522	522

Notes: The table above displays the results of regression $S_{it} = S_0 + \sigma * SE(S_{it}) + \gamma * SE(S_{it}) * X_t + \epsilon_{it}$, where S_{it} is the i-th estimate of size effect in study j and $SE(SE_{it})$ is the respective standard error. The X_t is either year of the publication of the study t, or an impact factor of the outlet, in which study t was published. In specification (1) OLS is used. Following specification (2) is panel data regression with between effects. The next specification (3) is estimated by WLS with precision used as weight. Similarly, the specification (4) is regressed. Here is the reciprocal of number of estimates reported per study used as a weight. The last specification (5) is the instrumental variables estimation. The reciprocal of the square root of the number of observations is employed as an instrument. The standard errors are clustered at the study level. In parentheses are reported the standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level.