LUND UNIVERSITY
School of Economics and Management

The Effect of Customer-base Diversification on Idiosyncratic Volatility

Written by Ronie Islam & Anton Törnblom

24th of May, 2017

Master's Programme in Finance
Supervisor: Charlotte Christiansen
Key Words: *Customer-base Concentration, Idiosyncratic Volatility, Diversification*

Abstract

The study sets out the objective to investigate how a firm's customer-base diversification affects its stock return idiosyncratic volatility. The research paper primarily focuses on the way that the distribution of sales impacts the firm specific risk that is unrelated to the risk of the market, and it does so through the use of both a geographical and an operating customer-base concentration measure. The previous literature and the underlying theory mainly propose that there is a positive relationship between customer-base concentration and idiosyncratic volatility. The findings of this study appear to be in line with that hypothesis at a first glance; however, the results lack significance when controlling for other firm specific factors. Preliminary testing that excludes the control variables suggests both a positive and significant effect between customer-base concentration and idiosyncratic volatility. The significance of this effect is absorbed by the control variables when introduced to the models indicating that the cause of an increase in idiosyncratic volatility is in fact attributable to other firm characteristics than customer-base concentration. The study is conducted on Swedish companies and uses a panel data setting with 60 different companies over a period of 10 years ranging from 2006 to 2015.

Table of Contents

1. Introduction....1
2. Theoretical Background....4
   2.1 Volatility & Modern Portfolio Theory....4
   2.2 Systematic and Idiosyncratic Volatility....4
   2.3 Idiosyncratic Volatility and Customer-base Concentration....5
   2.4 Foundation of Equity Valuation....7
3. Literature Review....8
   3.1 Applications & Drivers of Idiosyncratic Volatility....8
   3.2 Customer-base Concentration & Multinational Presence....9
4. Methodology....11
   4.1 Sample Selection....11
   4.2 Sources....11
   4.3 Customer-base concentration score (CBC)....12
   4.4 Idiosyncratic Volatility (IV)....13
   4.5 Control Variables....14
   4.6 Univariate t-tests....16
   4.7 Regression Specification....17
   4.8 Continent Specific Effects....18
5. Results....19
   5.1 Univariate t-tests....19
   5.2 Regressions....23
   5.3 Continent Specific Effects....26
6. Discussion....29
7. Conclusion....36
8. Appendix....37
9. References....43

1. Introduction

As companies and industries develop, the stylized fact of higher supply chain connectivity follows. The relationships between a supplier and its customers are successively enhanced and businesses become increasingly integrated. Though these tight connections between a supplier and its customers are essential for the success of the company, they potentially also create a situation in which the supplier becomes highly sensitive to the demand of fewer customers and thus more exposed to specific demand shocks. A company with a customer-base that is highly concentrated would be significantly affected by return shocks experienced by a specific customer since a company with fewer customers is increasingly dependent on each of its customers' demand. Using the same logic, a company with a less concentrated customer-base would not be as affected by a specific customer return shocks since the company less dependent on each of its customers' demand. This creates an incentive for companies to disperse its sales across the customer-base to get a diversification effect and reduce its volatility, similar to how portfolio diversification functions.

In this research paper, the effect of companies' customer-base structure, in terms of both geographical segments and operating segments, on the companies' idiosyncratic return volatility is investigated. The reason why idiosyncratic return volatility is looked at instead of the total return volatility is because of the interest to see customer-base diversification as a tool to deal with exposure to firm specific shocks that are unrelated to the overall market risk, which is in theory seen as diversifiable while shocks that are related to the overall market risk are seen as undiversifiable. Several questions are posed throughout the paper. Does a customer-base that is less concentrated come with a diversification effect that reduces companies' idiosyncratic volatility? Do firm characteristics such as size, leverage and dividends play a role? Do some geographical areas contribute more to idiosyncratic volatility than others? If so, could a customer-base with larger concentration still have a low idiosyncratic volatility if based in a certain region? Based on the underlying theory and the previous literature regarding the subject, which is discussed in detail in the coming sections, it is expected that this research will produce results indicating a positive relationship between customer-base concentration and idiosyncratic volatility.

Mihov and Naranjo (2017) find that firms with less concentrated customer-bases have lower idiosyncratic volatility and they also show that this relationship depends on what type of customers the company has. More specifically, Mihov and Naranjo (2017) provide evidence suggesting that there is a positive and significant relationship between customer-base and idiosyncratic volatility but that this effect applies only to a firm's corporate proportion of the customer-base and not to the firm's customers that are more stable such as governments. In their report, customer base concentration is measured in terms of sales to specific customers that account for at least 10% of total revenue. This paper extends this research but uses different measures of customer base concentration; two measures of customer base concentration are applied, one in terms of sales to specific geographical segments and another in terms of sales to specific operating segments. Using these two categorizations of sales instead of using customer specific sales ensures a certain degree of fundamental difference between customer-types and produces more comparable measures across firms. Since the demand correlations between customer types are not specified by Mihov and Naranjo (2017), a number of different customers could still have relatively highly correlated demands and thus in fact produce no diversification effect even though a low concentration score is produced. Therefore, categorizing the customers into geographical segments and operating segments ensures a fundamental difference between customer types and potentially avoids demand correlations that are not reflected in the concentration score. Dennis and Strickland (2009) also investigate the relationship between customer concentration and idiosyncratic volatility, and they too find a positive correlation the two variables. Similar to this paper they used reported operating segment as a measure for customer base concentration but they do so to explain the historical increase in aggregate idiosyncratic volatility rather than firm specific idiosyncratic volatility. Moreover, the following research further develops the understanding of companies' ability to reduce idiosyncratic volatility through their customer base structure. It primarily extends the research of Mihov and Naranjo (2017) and Dennis and Strickland (2009) by using different measures of customer-base concentration and idiosyncratic volatility, which emphasize the effect of both operating and geographical customer-base diversification on individual firm specific risk. The paper also contributes to the understanding of idiosyncratic volatility and the research in which idiosyncratic volatility is an important aspect.

To investigate the effect of customer base concentration, both in terms of operating segment and geographical segment, on idiosyncratic volatility in a structured and contextual manner this paper is divided up into several different sections: Section 2, which provides a theoretical background of the topic; Section 3, which presents the previous literature and research made in relation to the topic; Section 4, which describes the data and methods used in the research; Section 5 that provides representations and descriptions to the results of the study; Section 6 that interprets the results of the investigation and discusses them in relation to the hypothesis, the underlying theory and the findings of the previous literature; and Section 7 that summarizes the key points and concludes the findings of the research.

2. Theoretical Background

Presented below are the theoretical concepts on which the study is based on. The topics explained and discussed are: volatility in relation to modern portfolio theory; the differences between systematic and idiosyncratic volatility; customer-base concentration's effect on idiosyncratic volatility; and the foundations of equity valuation relating idiosyncratic volatility to uncertainty in cash flow streams.

2.1 Volatility & Modern Portfolio Theory

Volatility refers to the level of dispersion in stock returns. Volatility, commonly measured with variance or standard deviation, is a latent variable and can therefore only be estimated rather than calculated. Presented by Markowitz (1952), modern portfolio theory argues that an investor will only take on risk if it yields higher expected return, based on the assumption that investors are risk averse. The theory suggests that the characteristics of investments when considered individually as single investments are not as important as their characteristics in the context of the entire portfolio. Specifically, measures such as variance and correlation should be analyzed in terms of the effect on the entire portfolio's return and risk. This concept of portfolio risk is further explained by the portfolio return variance formula:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_i \sigma_j \rho_{ij}, \quad (1)$$

where, σ_p^2 is the variance of the portfolio, w_i weight of asset *i*, σ_i is the standard deviation of asset *i*, and ρ_{ij} is the correlation between asset *i* and asset *j*. In other words, modern portfolio theory suggests that the volatility of an investor's portfolio will decrease if the investor decides to invest in more assets given that the correlation between the assets is less than one. As the correlation between assets approaches one, the diversification effect dissipates.

2.2 Systematic and Idiosyncratic Volatility

Sharpe (1964) and Lintner (1965) argue that the risk of an asset can be specified into two different categories, market risk (also known as systematic risk and non-diversifiable risk) and asset specific risk (also known as idiosyncratic risk). The market risk refers to the risk associated to the entire market, such as government policy changes, natural disasters and other unexpected macroeconomic events. In contrast, idiosyncratic risk has little correlation to the market risk as it only includes the risk associated with the specific asset and does not affect the rest of the market.

In the capital asset pricing model presented by Sharpe (1964) and Lintner (1965), asset return is a function of time value of money and risk. The relationship is given by:

$$r_i = r_f + \beta_i(r_m - r_f), \quad (2)$$

where: r_f refers to the risk free return and describes the compensation given for the time value of money; $r_m - r_f$ refers to the market risk premium and describes the compensation given for the market risk.; and β_i is the factor of which the asset i is exposed to the market risk. Hence, the CAPM suggests that the only risk that an investor is compensated for is the systematic risk. This model is based on the same portfolio return variance formula presented by modern portfolio theory, which in essence shows that the idiosyncratic volatility can be diversified away by holding a portfolio of assets and should therefore not be priced.

As the total volatility equates to the sum of the systematic and idiosyncratic volatility, it is possible to derive the idiosyncratic portion by subtracting the systematic volatility, estimated using the CAPM, from the total volatility. This is equivalent to using the variance of the residuals taken from the CAPM model, which is the applied method to calculate the idiosyncratic volatility in this paper.

2.3 Idiosyncratic Volatility and Customer-base Concentration

Irvine and Pontiff (2009) presents a theoretical framework of how idiosyncratic volatility is dependent industry competitiveness through a Cournot equilibrium model. Mihov and Naranjo (2017) further adapt this framework and apply it to customer-base concentration. With their model, Irvine and Pontiff (2009) show that firm specific demand shocks cause a lower correlation between the firms' profits leading to more idiosyncratic volatility. Mihov and Naranjo (2017) then argue that this effect is amplified for suppliers with highly concentrated customer-bases due to the fact that firm specific customer demand shocks increasingly influence the total demand to the suppliers as the suppliers increase their customer-base concentration. In other words, the effect that the firm specific customer demand shocks have on the idiosyncratic volatility becomes more evident for supplier with a higher customer-base concentration.

The framework assumes an industry of two companies that experience zero-sum firm specific demand shocks, i.e. shifts in demand from one company to the other. Irvine and Pontiff (2009) derive that the profits of the two companies are given by:

$$\pi_1 = f(\theta, \mu, q_1, q_2, k, \omega_1) = (\theta - \mu - q_1 - kq_2)q_1 - \frac{\omega_1}{2} {q_1}^2 \qquad (3)$$

$$\pi_2 = f(\theta, \mu, q_1, q_2, k, \omega_2) = (\theta + \mu - q_2 - kq_1)q_2 - \frac{\omega_2}{2} {q_2}^2 \qquad (4)$$

where, π_i the profits of company i, θ describes an industry demand shock, μ is a zero-sum firm specific demand shock, k is an industry competitiveness parameter between 0 and 1 (a low value corresponds to low industry competition and a high value corresponds to high industry competition), ω_i is a stochastic variable that describes the input cost of company i and q_i describes its production quantity.[1] From this, the maximum profits given a Cournot equilibrium setting are derived:

$$\pi_{1,max} = f(\theta, \mu, k, \omega_1, \omega_2) = \frac{(\omega_1+2)\big((\theta-\pi)(\omega_2+2)-k(\theta+\pi)\big)^2}{2\big((2-k)(2+k)+2\omega_1+2\omega_2+\omega_1\omega_2\big)^2} \qquad (5)$$

$$\pi_{2,max} = f(\theta, \mu, k, \omega_1, \omega_2) = \frac{(\omega_1+2)\big((\theta+\pi)(\omega_2+2)-k(\theta-\pi)\big)^2}{2\big((2-k)(2+k)+2\omega_1+2\omega_2+\omega_1\omega_2\big)^2} \qquad (6)$$

where, $\pi_{1,max}$ is negatively related to μ and $\pi_{2,max}$ is positively related to μ, meaning that a firm specific shock affects the profits of company 1 negatively while the profits of company 2 are affected positively.[2] The model suggests that an industry shock, θ, contributes to a higher correlation between the firms' profits, while a firm specific shock, μ, as well as higher industry competitiveness, k, leads to a lower correlation. Conclusively, this demonstrates the concept that higher customer base-concentration, leading to a higher spread in μ, cause lower correlation between the firms' profits, which in turn leads to more idiosyncratic volatility according to Mihov and Naranjo (2017).

---

[1] These profit functions are based on the following cost and price functions:

$C_1 = \frac{\omega_1}{2} {q_1}^2, \quad C_2 = \frac{\omega_2}{2} {q_2}^2, \quad p_1 = \big(\theta - \mu - q_1 - kq_2\big)q_1, \quad p_2 = \big(\theta + \mu - q_2 - kq_1\big)q_2.$

[2] The optimal production quantities in Cournot equilibrium are given by:

$q_1^* = \frac{(\theta-\mu)(\omega_2+2)-k(\theta+\pi)}{(2-k)(2+k)+2\omega_1+2\omega_2+\omega_1\omega_2} \quad q_2^* = \frac{(\theta+\mu)(\omega_1+2)-k(\theta-\pi)}{(2-k)(2+k)+2\omega_1+2\omega_2+\omega_1\omega_2}$

2.4 Foundation of Equity Valuation

The economic intuition behind the potential relationship of customer-base concentration on stock return volatility could be explained through a cash-flow perspective since, according to the foundations of equity valuation, stock prices reflect the present value of future cash-flow streams given a rational market (Koller, Goedhart and Wessels, 2015).[3] Since the future cash-flows are often unknown, today's value of the stock is based on the expected future cash-flows:

$$P_0 = \sum_{i=1}^{\infty} \frac{E_0(CF_i)}{(1+k)^i}, \tag{7}$$

where P_0 is the price of the stock as of today, E_0 is the expectation as of today, CF_i is the cash-flow at time *i,* and k is the proper discount rate. Changes in the stock price are by this logic based on changes of expected future cash-flows. This means that the volatility of stock returns comes from the volatility of unexpected cash-flows. More specifically, volatility in the stock prices, and correspondingly in the stock returns, are determined by unexpected changes to the cash-flow stream and not by the cash-flows themselves. The volatility of the cash-flow streams could be influenced by the customer concentration structure as a more diverse customer-base has a potential cash-flow smoothening effect and thus also reduce unexpected cash-flow changes and that results in lower volatility of the stock prices.

[3] The other factor that drives the stock returns besides the future cash flow streams is the discount rate; however, T. Vuolteenaho (2002) finds that only a third of the value is attributable to the discount rate.

3. Literature Review

The following research paper extends the literature investigating the relationship between customer-base concentration and idiosyncratic volatility. As mentioned, two articles that have previously investigated the specific relationship between customer-base concentration and idiosyncratic volatility are Mihov and Naranjo (2017) and Dennis and Strickland (2009). Even though there are few others that have investigated this relationship, the literature that examines the factors affecting idiosyncratic volatility and the literature that covers the topic customer-base structure are both vast.

3.1 Applications & Drivers of Idiosyncratic Volatility

This research paper contributes to the understanding idiosyncratic volatility, which is applicable in a number of different fields including: pricing of idiosyncratic volatility resulting from under-diversification; understanding the total volatility of stock prices which is an important element in pricing models and valuing options and derivatives; and idiosyncratic volatility's effect on corporate policies such as cash holdings, dividend policy and financial leverage. In addition, the paper also contributes to the discussion of what is driving idiosyncratic volatility.

Goyal and Santa-Clara (2003) comes to the conclusions that idiosyncratic risk matters in the sense that it corresponds to a significant portion of total asset return volatility and thus affects the prices. Though there are extensions to the CAPM and other models that account for the price of idiosyncratic risk, the findings of Goyal and Santa-Clara (2003) contradict the basis on which many economic models rely on, including the original CAPM where only systematic risk affects the asset prices. Goetzmann and Kumar (2008) and Polkovnichenko (2005) provide evidence that individual investors do not tend to have fully diversified portfolios and this is a potential economic explanation behind the prices attributed to idiosyncratic volatility. The extensions and models that have been developed to account for idiosyncratic volatility, including Malkiel and Xu (2001), are therefore often based on under-diversified portfolios.

As total volatility partly consists of idiosyncratic volatility, the idiosyncratic volatility becomes important in valuing options and derivatives. For example, the option pricing model presented by Black and Scholes (1973) uses the total volatility to derive the option prices and therefore an understanding of how the components of the total volatility behave is important. In other words, both the systematic volatility and the idiosyncratic volatility are vital in estimating accurate asset prices.

Idiosyncratic volatility is an important element in the topic of corporate policies. Chay and Suh (2009) show that corporate payout policy is heavily influenced by cash-flow uncertainty; high stock return volatility, which is used as a proxy for cash-flow uncertainty, has a negative impact on the amount of dividends paid out. The uncertainty in cash-flow streams also has an influence on companies' cash holdings, where firms' cash ratios tend to increase as cash-flow streams become more uncertain (Bates, Stultz and Kahle, 2009). As documented by Bartram, Brown and Waller (2015) and Markarian and Gill-de-Albornoz (2012) other corporate policy decisions influenced by idiosyncratic volatility includes leverage and income smoothening.

The paper also adds to the discussion of what is driving idiosyncratic volatility. In addition to the previously mentioned concept of customer structure being one of the drivers of idiosyncratic volatility there is research suggesting that market composition and ownership structure effects idiosyncratic volatility as well. Gaspar and Massa (2006) put forward evidence indicating that firms with high market power and firms operating in industries with fewer competitors tend to have lower idiosyncratic volatility. Malkiel and Xu (2003) argues that there is a strong positive relationship between idiosyncratic volatility and growth opportunities as well as ownership structure; firms being owned by institutional investors tend to have higher idiosyncratic volatility.

3.2 Customer-base Concentration & Multinational Presence

Furthermore, this research is also relevant to the topic on the consequences of customer base structure and multinational presence. Patatoukas (2012) shows that firms with large customers in relation to their sales, i.e. with a concentrated customer base, appears to be more efficient as it reduces operating expenses, giving higher profitability and stock returns. There is also research indicating that multinational companies exhibit higher stock returns than non-multinational companies. Fillat and Garetto (2015) argue that although companies with a multinational customer base potentially become more geographically diversified, the exposure to cash-flow risk is higher as multinational firms are hesitant to withdraw from markets abroad due to the sunk costs from entering the markets. There is also previous research that investigates the idea of having a customer base that is concentrated in emerging markets and how it influences risk. Angelidis (2010) proposes that the dynamics of idiosyncratic volatility potentially differs depending on geographical region, and in particular finds evidence suggesting that the idiosyncratic volatility of firms in emerging markets follow different trends than that of firms in developed markets.

4. Methodology

Below follows the sample selection process, the data sources, the computation of the customer-base concentration measures, the derivation of the idiosyncratic volatility, the specifications and motivations for each of the control variables, the methodology of the univariate t-tests, the specification methods of the regression models and the methods used to investigate continent specific effects.

4.1 Sample Selection

The first step of the selection process is to list Swedish companies with at least 10 years of trading history sorted by largest to smallest in terms of trading turnover value. According to the Swedish accounting council (Redovisningsrådet), Swedish companies are recommended to report segments corresponding to at least 10% of the total sales. [4] Though only the segments that correspond to at least 10% are recommended, it is common that all segments, regardless of size, are reported by the companies. Out of the first 100 companies from the list 60 of these reported sales for all 10 periods between 2006 and 2015 and for all segments regardless of size in terms of both geographical and operating segments. The 40 companies that were excluded due to insufficient data could potentially raise an issue of sample selection bias. However, if a company did or did not report sufficient data seem to be arbitrary as no evident systematic pattern among the 40 companies was observable. The complete list of the 60 companies investigated is found in Appendix, table 1.

4.2 Sources

All data needed to derive the customer-base concentration scores, including sales by geographical segment and sales by operating segment, is extracted from Bloomberg. The sales data provided by Bloomberg are consistent with the annual reports of the companies, meaning that the data reflects the sales reporting of the companies themselves. Thus, missing data points in the Bloomberg database are complemented

[4] Information about segment specific sales, defined both by geographical segment and operating segment, are recommended to be disclosed given that the sum of the internal and external sales from the segment corresponds to 10% or more of the total internal and external sales of the company (Redovisningsrådet, 2002).

using the annual reports of the companies. The data needed to derive the idiosyncratic volatility, with exception to the risk free rate (i.e. including the stock prices and the proxy for the market return), is extracted from Thomson Reuters. The proxy for the risk free rate is instead extracted directly from database of Riksbanken, the central bank of Sweden. Thomson Reuters is also used to collect data for all control variables used in the regressions.

4.3 Customer-base concentration score (CBC)

In order to obtain comparable geographical segments, the sales data for each company is categorized in terms of the following continents: *Africa & Middle East*, *Asia & Oceanic*, *Europe*, *North America (NAFTA)* and *South & Central America*.[5]

The operating segments are defined by how they are reported by the companies themselves and no recategorization is made. Companies are recommended to report operating segments according to the standards of IFRS8, which requires firms to disclose operating segment information based on internal management reports (Ernst & Young, 2009).

To measure the level of concentration, a Herfindahl-Hirschman index (HHI) is used which is defined similar to Patatoukas (2012), i.e. by summing the squared shares of each segment. The HHI scores, which can range between 0 and 10’000 are then standardized yielding a customer-base concentration score between 0 and 1, where a customer-base concentration score of 0 corresponds to complete operating or geographical diversification and customer-base concentration score of 1 corresponds to no operating or geographical diversification. The equations to the CBC measures are given by:

$$CBC_{Geo,t} = \frac{HHI_{Geo,t} - 2000}{8000} = \frac{\sum_{i=1}^{5} S_{i,t}^2 - 2000}{8000} \qquad (8)$$

$$CBC_{Op,t} = \frac{HHI_{Op,t}}{10000} = \frac{\sum_{i=1}^{n} S_i^2}{10000} \qquad (9)$$

[5] If country specific sales were reported, the United Nations Statistics Division (2017) was used to assign sales of countries to corresponding continents.

$S_{i,t}$ corresponds to the sales of segment *i* expressed as a percentage of total sales for year *t*. Note that in order to get standardized values that range between 0 and 1, the HHI scores are subtracted by 2000 and then divided by 8000 for the geographical CBC since the maximum number of segments is 5 meaning that the smallest possible HHI score is 2000.

Table 1: CBC Descriptive Statistics

Table 1 contains descriptive statistics including the number of observations, mean, median and standard deviation of the CBC measure for each year and in total.

	N	Mean CBC_{Op}	Median CBC_{Op}	Std. Dev. CBC_{Op}	Mean CBC_{Geo}	Median CBC_{Geo}	Std. Dev. CBC_{Geo}
2006	60	0.59	0.51	0.27	0.73	0.87	0.31
2007	60	0.56	0.51	0.27	0.71	0.84	0.32
2008	60	0.56	0.50	0.27	0.70	0.84	0.33
2009	60	0.55	0.46	0.27	0.68	0.83	0.33
2010	60	0.56	0.49	0.28	0.68	0.80	0.34
2011	60	0.57	0.51	0.28	0.67	0.82	0.35
2012	60	0.57	0.49	0.28	0.66	0.76	0.35
2013	60	0.54	0.50	0.26	0.65	0.74	0.35
2014	60	0.54	0.44	0.26	0.65	0.72	0.35
2015	60	0.53	0.45	0.26	0.64	0.67	0.36
Total	600	0.56	0.50	0.27	0.68	0.79	0.34

4.4 Idiosyncratic Volatility (IV)

Since idiosyncratic volatility, as well as total volatility, is a latent variable it has to be estimated. The method used in order to estimate the idiosyncratic volatility is similar to that of Hoberg and Prabhala (2009); it is measured as the variance of the residuals derived from the CAPM, where OMXS30 is used to estimate of the market returns and the SE 3M (Swedish three month treasury bill) is used to estimate the risk free rate. Continuously compounded returns, derived from using the log-differencing method on the closing prices, are used. The firm specific betas are derived by running an OLS-regression (using the CAPM) on the previous 250 trading days and these betas are reestimated at the end of each year. As the skewness of idiosyncratic volatility potentially has a negative influence on the statistical inferences, a natural log-transformation is made on the idiosyncratic volatility values (Gaspar and Massa, 2006).

For a three-dimensional visual representation of the idiosyncratic volatility by firm and year see Appendix, figure 1. To see the difference in the idiosyncratic volatility, IV and the natural log of the idiosyncratic volatility, ln(IV), see Appendix, figure 2 and Appendix, figure 3.

Table 2: IV Descriptive Statistics

Table 2 contains descriptive statistics including the number of observations, mean, median, standard deviation, skewness and kurtosis of ln(IV) for each year and in total.

	N	Mean $ln(IV)$	Median $ln(IV)$	Std. Dev. $ln(IV)$	Skewness $ln(IV)$	Kurtosis $ln(IV)$
2006	60	-8.08	-8.21	0.69	0.82	4.12
2007	60	-8.03	-8.04	0.61	0.46	4.11
2008	60	-7.32	-7.30	0.46	0.36	3.02
2009	60	-7.55	-7.63	0.66	0.89	4.17
2010	60	-8.29	-8.36	0.79	1.39	6.47
2011	60	-8.17	-8.19	0.73	0.76	4.04
2012	60	-8.41	-8.59	0.89	1.49	6.54
2013	60	-8.55	-8.59	0.85	1.39	6.00
2014	60	-8.59	-8.64	0.88	1.04	4.62
2015	60	-8.30	-8.51	0.87	1.65	5.90
Total	600	-8.13	-8.19	0.85	0.62	3.64

4.5 Control Variables

The control variables have been selected based on previous studies and theories to strictly assess the effect of customer-base consecration on idiosyncratic volatility. Below follow definitions, rationale, and corresponding references for all the considered control variables.

Age – Age is defined as the natural log of the number of years that the company has been present on the Stockholm stock exchange. The variable is used to control for investor uncertainty. The natural log is used to reflect the fact that one year of age has more influence on investor uncertainty when it comes to young firms compared old firms (Pástor and Veronesi, 2003).

Analysts – This control variable is defined as the number of analysts following the stock. Morck, Yeung and Yu (2000) find that there is a relationship between idiosyncratic volatility and how well company information is publically available to investors. Similarly to Mihov and Naranjo (2017), the number of analysts is used as a proxy for the availability of company information.

Beta – Beta refers to the stocks' substitutability, i.e. if there are any close substitutes to the stock. Similar to Gaspar and Massa (2006), the CAPM beta coefficient is used as a proxy to measure the substitutability. As previously mentioned, the firm specific betas are derived by running an OLS-regression (using the CAPM) on the previous 250 trading days where the betas are reestimated at the end of each year. As explained by Wurgler and Zhuravskaya (2002), stocks that are more easily substitutable have lower idiosyncratic risk and hence this variable is controlled for.

Div – This control variable is a dummy variable, which takes on the value 1 (and otherwise 0) if the company pays out dividend for the corresponding year. This variable is controlled for since firms that pay no dividends tend to have more volatile returns (Pástor and Veronesi, 2003).

Lev – Lev is defined as the end of year book value of total debt over the end of year book value of total assets, i.e. describing the leverage for each firm. The variable controls for the risk induced by increasing the level of indebtedness (Black, 1976).

M/B – M/B is defined as the end of year market value of equity over the end of year book value of equity for each firm. The variable controls for growth opportunities as this may affect the volatility of the returns (Mihov and Naranjo, 2017).

OMXS30 – This control variable is a dummy variable, which takes on the value 1 (and otherwise 0) if the company is a part of the OMXS30 index during the corresponding year, i.e. if the company is among the most traded stocks in terms of value on the Stockholm stock exchange. This variable is controlled for in order to identify if the results are specific to large and mature firms that are commonly traded (Gaspar and Massa, 2006).

Price – The price is defined as the closing price at the end of the each year. The variable controls for microstructure noise on the volatility estimates, which can potentially arise for low price stocks (Gaspar and Massa, 2006).

ROA – ROA is defined as the ratio between the operating profits and total assets, i.e. the return on assets. This variable controls for the level of profitability for each firm (Gaspar and Massa, 2006).

Size – Size is defined by the end of year market capitalization expressed in millions for each firm. The variable controls for potential size effects (Mihov and Naranjo, 2017).

TO – The TO is defined as the average monthly number of shares traded divided by the total shares outstanding, using the last month for each year. This variable controls for the positive correlation between return volatility and volume (Gaspar and Massa, 2006).

Table 3: Control Variable Descriptive Statistics

Table 3 contains descriptive statistics including mean, median, maximum, minimum, standard deviation, skewness, kurtosis and number of observations of all control variables.

	Age	Analysts	Beta	Div	Lev	M/B	OMXS30	Price	ROA	Size	TO
Mean	2.97	11.99	1.81	0.85	0.23	11.16	0.33	103.77	0.06	51866.09	0.01
Median	2.83	9.00	1.76	1.00	0.23	2.20	0.00	81.83	0.05	12833.26	0.01
Max	4.96	44.00	4.45	1.00	0.80	4848.96	1.00	726.58	0.83	726467.50	0.18
Min	0.69	0.00	-0.73	0.00	0.00	0.07	0.00	0.37	-0.62	61.16	0.00
Std. Dev.	0.83	9.81	0.77	0.36	0.18	197.86	0.47	89.72	0.12	98874.23	0.01
Skewness	0.39	0.87	0.11	-1.96	0.45	24.42	0.71	2.33	0.14	3.12	7.67
Kurtosis	3.03	2.79	2.73	4.84	2.37	597.66	1.50	11.77	15.24	14.40	86.11
N	600	600	600	600	600	600	600	600	600	600	600

4.6 Univariate t-tests

In order to get an idea to whether or not there are any statistically significant differences in idiosyncratic volatility between companies with high CBC and low CBC, two different types of univariate t-tests are conducted. The first univariate t-test is conducted by looking at the difference in the idiosyncratic volatility between the companies with an above median CBC and the companies with a below median CBC. The second univariate t-test looks at the difference in idiosyncratic volatility between companies with a CBC in the first quartile and the companies with a CBC in the fourth quartile. The tests are done for each of the ten separate years as well as for the total sample period, using the mean log idiosyncratic volatility for the firms.

4.7 Regression Specification

In order to investigate the relationship between idiosyncratic volatility and customer base concentration panel regressions are run using *EViews 9* with a total of 600 observations (60 cross-sectional units and 10 periods). Regressions are run for both CBC measures separately and combined, as well as with and without the applicable control variables.

To obtain the right specifications for the models, the first test conducted is a test for potential multicollinearity between the explanatory variables. The problems of having near multicollinearity between the explanatory variables include obtaining misspecified beta-coefficients, which increase the chances of making incorrect inferences. In order to identify near multicollinearity, a correlation matrix is used and correlations that exceed 0,8 are considered to be problematic.[6]

Secondly, non-stationarity tests are conducted for all applicable variables. A potential problem of having non-stationary variables is that it might create spurious relationships between variables. Furthermore, inferences could be distorted since distribution assumptions are no longer valid. The test used in this paper in order to identify non-stationary variables is the augmented Dickey-Fuller test.

Thirdly, it is determined if pooled regressions can be applied or if there is a need to use fixed effects models, alternatively random effects models. This is done by testing for

---

[6] Using a correlation of 0,8 as a threshold to identify near multicollinearity is a practice proposed Brooks (2014).

heterogeneity in both the cross-sectional and time dimensions using the redundant fixed effects likelihood ratio test. If heterogeneity is found, random effects models are preferred to fixed effects models since the interpretation of the numerical values stay the same and since the random effects models preserve a high degree of freedom giving smaller standard errors. However, in contrast to the fixed effects models where potential endogeneity problems are eliminated, endogeneity might still prevail in random effects models. Therefore, the Hausman specification test is conducted on all random effects models. When endogeneity is found, fixed effects models are applied.

Moreover, White's robust standard errors are used for all models to account for potential heteroscedasticity. The need for robust standard errors is due to the problems that occur when the variance of the error terms differ across variables. For example, heteroscedasticity could lead to erroneous standard errors, which increase the chances of making incorrect inferences.

4.8 Continent Specific Effects

To add further to the investigation, continent specific effects are examined through additional regressions. This is mainly done in order to see if sales in certain continents influence idiosyncratic volatility in which case continent specific effects is something that needs to be accounted for when investigating the relationship between customer-base concentration and idiosyncratic volatility. In order to do this, the sales percentages in each of the earlier defined continents, *Africa & Middle East*, *Asia & Oceanic*, *Europe*, *North America (NAFTA)* and *South & Central America* are regressed against the idiosyncratic volatility together with the control variables specified earlier. Since the sales percentages variables are complementary, i.e. always add up to 1, they are perfectly multicollinear and can therefore not be regressed simultaneously. Hence, each continent variable is regressed separately creating five different regressions. A sixth regression is also made with another variables defined by the percentage sales in *Europe* and *NAFTA* combined; this variable aims to reflect the amount of sales located in developed markets in comparison to emerging markets (*Africa & Middle East*, *Asia & Oceanic* and *South & Central America*). The regression specifications are determined with the same procedure as shown in Section 4.7.

5. Results

Below follows the results of: univariate t-tests, which compares the average idiosyncratic volatility for companies with different levels of customer-base concentration; the regression specifications and the regressions models for each CBC measure regressed independently and combined, with and with control variables; and the regressions investigating potential continent specific effects.

5.1 Univariate t-tests

Univariate t-tests are conducted in order to preliminarily examine whether or not there is a difference in average idiosyncratic volatility between companies with high CBC and low CBC.

Table 4: Univariate t-test – Geographical CBC Halves Comparison

*Table 4 shows the statistics for the variables involved in the first univariate t-test based on the geographical measure of customer-base concentration. Column (1) describes the median CBC. Column (2) and (3) describe the corresponding mean in ln(IV) of the below median CBC firms, respectively above median CBC firms. Column (4) describes the difference between column (2) and (3). Column (5) reports the t-statistic given a two-sided t-test on whether column (4) is significantly different from zero. T-statistics with the notation of a *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.*

	(1) Median CBC_{Geo}	(2) Mean $ln(IV)$ H1	(3) Mean $ln(IV)$ H2	(4) Difference	(5) T-Stat
2006	0.87	-8.15	-8.00	0.15	0.84
2007	0.84	-8.10	-7.96	0.14	0.89
2008	0.84	-7.38	-7.27	0.11	0.87
2009	0.83	-7.67	-7.43	0.24	1.42
2010	0.80	-8.48	-8.09	0.38	1.92*
2011	0.82	-8.29	-8.05	0.24	1.30
2012	0.76	-8.48	-8.33	0.14	0.61
2013	0.74	-8.63	-8.46	0.17	0.78
2014	0.72	-8.56	-8.62	-0.06	-0.25
2015	0.67	-8.18	-8.42	-0.25	-1.09
Total	0.79	-8.19	-8.06	0.13	1.85*

The differences in average $ln(IV)$ between firms with low geographical CBC and high geographical CBC, presented in column (4) of Table 4, show ambiguous results as firms with higher customer-base concentration have higher average idiosyncratic volatility for the first eight years but lower for the last two years. This difference is statistically significant at the 10% level for the total sample period; however, almost none of the differences are significant when each year is looked at separately.

Table 5: Univariate t-test – Operating CBC Halves Comparison

Table 5 shows the statistics for the variables involved in the first univariate t-test based on the operating measure of customer-base concentration. It is structured in the same way as Table 4.

	(1) Median CBC_{Op}	(2) Mean $ln(IV)$ H1	(3) Mean $ln(IV)$ H2	(4) Difference	(5) T-Stat
2006	0.51	-8.27	-7.89	0.38	2.19**
2007	0.51	-8.19	-7.87	0.33	2.15**
2008	0.50	-7.38	-7.26	0.12	1.02
2009	0.46	-7.57	-7.54	0.03	0.16
2010	0.49	-8.53	-8.04	0.49	2.48**
2011	0.51	-8.36	-7.97	0.39	2.12**
2012	0.49	-8.57	-8.24	0.32	1.41
2013	0.50	-8.70	-8.39	0.31	1.42
2014	0.44	-8.69	-8.49	0.19	0.85
2015	0.45	-8.39	-8.21	0.19	0.83
Total	0.50	-8.27	-7.99	0.27	4.02***

The differences in average $ln(IV)$ between firms with low operating CBC and high operating CBC, presented in column (4) of Table 5, show that the idiosyncratic volatility is lower for firms with below median customer-base concentration than for firms with above median customer-base concentration for all the investigated years. Even though this difference is not statistically significant for all the separate years independently, it is significant for the total sample period at the 1% level.

The results in Table 4 and Table 5 suggest that firms with higher customer-base concentration, both based on the geographical and operational measure, experience higher stock return idiosyncratic volatility. When comparing the results between the two tables it is noticeable that the differences in idiosyncratic volatility, shown in column (4) of each table, are larger in absolute values as well as more significant when using the operational CBC measure. These results suggest that further investigation of the relationship is relevant and therefore univariate t-tests are also conducted using quartiles, which ensure that there is a larger difference in CBC between the two groups compared.

Table 6: Univariate t-test – Geographical CBC Quartiles Comparison

Table 6 show the statistics for the variables involved in the second univariate t-test based on the geographical measure of customer-base concentration. Column (1), column (2), column (3), and column (4) describe the mean ln(IV) of the firms with CBC in the first, second, third and fourth quartile respectively.[7] *Column (5) describes the difference between column (4) and (1). Column (6) reports the t-statistic given a two-sided t-test on whether column (5) is significantly different from zero. T-statistics with the notation of a *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.*

	(1) Mean $ln(IV)$ Q1	(2) Mean $ln(IV)$ Q2	(3) Mean $ln(IV)$ Q3	(4) Mean $ln(IV)$ Q4	(5) Difference	(6) T-Stat
2006	-8.34	-7.97	-7.45	-8.21	0.14	0.79
2007	-8.24	-7.95	-7.49	-8.10	0.14	0.88
2008	-7.52	-7.23	-7.07	-7.34	0.18	1.52
2009	-7.72	-7.62	-7.07	-7.57	0.16	0.84
2010	-8.47	-8.49	-7.77	-8.23	0.23	0.93
2011	-8.40	-8.18	-7.59	-8.24	0.16	0.74
2012	-8.64	-8.31	-7.81	-8.56	0.09	0.36
2013	-8.71	-8.55	-7.88	-8.75	-0.04	-0.18
2014	-8.77	-8.35	-8.22	-8.79	-0.02	-0.06
2015	-8.50	-7.85	-8.08	-8.57	-0.07	-0.31
Total	-8.33	-8.05	-7.66	-8.23	0.11	1.35

When observing the results of the first three quartiles, presented in column (1) through (3) of Table 6, the idiosyncratic volatility appears to increase as the customer-base concentration increases. Worth investigating further is however the fact that the idiosyncratic volatility seems to be lower in the fourth quartile, consisting of the companies with a CBC-score of 1, compared to third quartile. Nonetheless, the idiosyncratic volatility is higher in the fourth quartile than in the first quartile in total and for most of the years, which is indicated by column 5. This suggests that there some

[7] Note that firms with a CBC-score of 1 (which is more than 25% of the firms for the geographical measure) are all included in the fourth quartile.

difference in idiosyncratic volatility between firms with high and low customer-base concentration, yet none of the differences are statistically significant. As the results are significant when firms with below median CBC and above median CBC were compared, shown in Table 4, these results perhaps suggest that firms with a CBC of 1 behave differently.

Table 7: Univariate t-test – Operating CBC Quartiles Comparison

Table 7 show the statistics for the variables involved in the second univariate t-test based on the operating measure of customer-base concentration. The table is structured in the same manner as Table 6.

	(1) Mean $ln(IV)$ Q1	(2) Mean $ln(IV)$ Q2	(3) Mean $ln(IV)$ Q3	(4) Mean $ln(IV)$ Q4	(5) Difference	(6) T-Stat
2006	-8.40	-8.14	-8.11	-7.67	0.72	2.91***
2007	-8.29	-8.10	-7.92	-7.81	0.48	2.08**
2008	-7.42	-7.35	-7.37	-7.15	0.27	1.60
2009	-7.65	-7.48	-7.50	-7.57	0.08	0.37
2010	-8.57	-8.49	-8.10	-7.99	0.59	2.10**
2011	-8.49	-8.23	-8.12	-7.83	0.66	2.53**
2012	-8.62	-8.51	-8.30	-8.19	0.44	1.11
2013	-8.67	-8.73	-8.21	-8.57	0.10	0.28
2014	-8.63	-8.74	-8.35	-8.63	0.00	-0.01
2015	-8.42	-8.36	-8.07	-8.34	0.08	0.22
Total	-8.32	-8.21	-8.01	-7.98	0.34	3.26***

When observing the results of the fourth quartiles, presented by column (1) through (4) of Table 7, the idiosyncratic volatility appears to increases with the operating customer-base concentration measure. For most of the separate years as well as for the whole sample period, the idiosyncratic volatility increases successively for each quartile. The difference between the first and the forth quartile, shown in column (5), is statistically significant at the 1% level for the total sample period. This is consistent with results of Table 5, where firms with below median CBC and above median CBC are compared since firms with below median CBC-score experienced significantly lower idiosyncratic volatility on average compared to firms with above median CBC in that test.

Both Table 5 and Table 7, in which the operational CBC measure is used, show results in line with the hypothesis that higher customer-base concentration increases stock return idiosyncratic volatility. The results of the geographical CBC measure, presented in Table 4 and Table 6, do not indicate as strong of a relationship. Given by the fact that the firms with high geographical CBC tends to have higher average idiosyncratic volatility there are still some indications that CBC makes a difference also for the geographical measure; but, in contrast to the results of the operating CBC measure, the difference is not statistically significant.

5.2 Regressions

Three different regression models are desired in order to fully investigate the relationship between customer-base concentration and idiosyncratic volatility. Each customer base-concentration measure is regressed separately as this demonstrates their independent effects on idiosyncratic volatility. Since the CBC measures are rather different as one describes the diversification in terms of geographical segment sales and the other describes the diversification in terms of operating segment sales, a regression model using both CBC measures included simultaneously is also used. Regressing with both measures simultaneously is made possible by their low correlation indicating no multicolinearity problem (see Appendix, Table 2). In other words, the third regression is applied since it shows how the idiosyncratic volatility behaves when accounting for two different types of diversifications simultaneously. Also note that the three regressions are run both with and without the control variables in order to obtain a complete understanding of the dynamics.

In order to obtain the correct specifications of the regression models a number of tests are performed and can be found in the Appendix. As mentioned previously, statistical checks are done for multicollinearity, non-stationarity, heterogeneity and endogeneity. Potential heteroscedasticity is also accounted for by the use of White's robust standard errors.

As there is a correlation of more than 0,8 between the *OMXS30* and *Analysts* (see Appendix, Table 2), this suggest that there might be a multicollinearity issue present between the two variables. A solution to this multicollinearity issue is to drop one of the two correlated variables. In contrast to the *Analysts* variable, the *OMXS30* variable is a dummy variable and therefore it contains less information. Furthermore, the effect that *OMXS30* controls for is already somewhat accounted for by other control variables including *Size*, *Age* and *TO*. Hence, the *OMXS30* control variable is dropped to avoid multicollinearity issues.

The unit root tests (see Appendix, Table 3) indicate that any suspicion of non-stationary variables that could cause spurious relationships can be rejected and that all variables can be used in levels with no differencing needed.

The heterogeneity tests (see Appendix, Table 4, Table 5, Table 6), indicate that there is significant heterogeneity in both dimensions for all thee regressions. These results eliminate the option of using pooled regressions since this would not capture all the period and firm information in the data. The results instead suggest that there is a need for using either fixed effects or random effects in the regression models. As the Hausman test (see Appendix, Table 7) illustrates evidence for endogeneity when applying random effects to the models, this means that the most statistically sound specification to use for all three regressions is a regression model with fixed effects in both the cross-sectional dimension as well as in the time dimension. The three models can now be specified using the following equations:

$Regression_1$:

$$ln(IV)_{i,t} = \alpha + \beta_{geo} CBC_{geo,i,t} + \sum_{x=1}^{10} \beta_x CV_{x,i,t} + \lambda_t + \mu_i + v_{it}, \tag{10}$$

$Regression_2$:

$$ln(IV)_{i,t} = \alpha + \beta_{op} CBC_{op,i,t} + \sum_{x=1}^{10} \beta_x CV_{x,i,t} + \lambda_t + \mu_i + v_{it}, \tag{11}$$

$Regression_3$:

$$ln(IV)_{i,t} = \alpha + \beta_{geo} CBC_{geo,i,t} + \beta_{op} CBC_{op,i,t} + \sum_{x=1}^{10} \beta_x CV_{x,i,t} + \lambda_t + \mu_i + v_{it}, \tag{12}$$

where: α is the panel regression intercept; the betas are coefficients to the corresponding explanatory variables; $CV_{x,i,t}$ represents control variable *x* (of the in total 10 control variables), for firm *i* at time *t*; λ_t is a time-varying intercept that captures all of the variables that affect $ln(IV)_{i,t}$ and that vary over time but are constant cross-sectionally; μ_i is a variable that is constant over time but varies cross-sectionally in order to capture firm-specific effect; and v_{it} is the remainder disturbance term that varies both in time and cross-sectionally and that captures everything that is left unexplained about $ln(IV)_{i,t}$. To clarify, λ_t and μ_i are the fixed effects variables used to capture the heterogeneity in the time dimension and the cross-sectional dimension respectively.

Table 8: Regressions

*Table 8 shows all three regression models using fixed effects in both dimensions. The three models are displayed with and without the control variables included. T-statistics with the notation of a *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.*

	(1) $Regression_1$ (excl. Control Variables)	(2) $Regression_1$	(3) $Regression_2$ (excl. Control Variables)	(4) $Regression_2$	(5) $Regression_3$ (excl. Control Variables)	(6) $Regression_3$
α	-8.3500	-7.7036	-8.4318	-7.9584	-8.4910	-7.8720
T-Stat	*-122.21******	*-19.80******	*-109.32******	*-19.17******	*-104.27******	*-18.64******
CBC_{Geo}	0.3275	-0.1522			0.1664	-0.1836
T-Stat	*3.48******	*-0.66*			*1.55*	*-0.79*
CBC_{Op}			0.5452	0.2320	0.4491	0.2512
T-Stat			*3.90******	*1.07*	*2.83******	*1.17*
Age		0.0486		0.0515		0.0550
T-Stat		*0.44*		*0.47*		*0.50*
Analysts		-0.0068		-0.0070		-0.0062
T-Stat		*-0.87*		*-0.91*		*-0.79*
Beta		-0.0038		-0.0066		-0.0057
T-Stat		*-0.10*		*-0.17*		*-0.15*
Div		-0.3449		-0.3183		-0.3199
T-Stat		*-2.16****		*-1.98****		*-1.98****
Lev		0.1623		0.1589		0.1993
T-Stat		*0.63*		*0.63*		*0.77*
M/B		0.0000		0.0000		0.0000
T-Stat		*-0.14*		*0.07*		*-0.21*
Price		-0.0012		-0.0013		-0.0013
T-Stat		*-2.56****		*-2.70******		*-2.64******
ROA		0.3977		0.4304		0.4388
T-Stat		*1.41*		*1.54*		*1.55*
Size		0.0000		0.0000		0.0000
T-Stat		*0.67*		*0.77*		*0.66*
TO		-5.3920		-5.2773		-5.5169
T-Stat		*-1.41*		*-1.39*		*-1.45*
R^2	0,7957	0,8179	0,7961	0,8183	0,7963	0,8186
Adj. R^2	0,7691	0,7903	0,7695	0,7907	0,7694	0,7906

Consistent with the results from the univariate t-tests previously conducted, it can be noted that the sign of all CBC measures are positive, i.e. as they increase so does $ln(IV)$ for all regressions when excluding the control variables. These regressions, shown in column (1), (3) and (5) of Table 8, also indicate that the operating CBC measure has a larger effect on $ln(IV)$ than the geographical CBC measure. Furthermore, the CBC measures are significant for $Regression_1$ and $Regression_2$; however, the geographical CBC measure is not statistically significant in $Regression_3$ where both measures are regressed simultaneously shown in column (5) of Table 8.

When adding the control variables to the regressions the effect of both CBC measures on idiosyncratic volatility decreases in all three regression models, which can be seen by the coefficients in column (2), (4) and (6) of Table 8. As for the control variables themselves few of them have coefficients that are significantly different from zero as only *Div* and *Price* are significant at a 5% level. Tough most are insignificant, their inclusion do have an effect on the CBC coefficients. After the inclusion of all control variables the coefficient for CBC_{Geo} changes sign in the regressions, meaning that higher customer-base concentration now decreases idiosyncratic volatility based on these models. Moreover, the CBC measures are no longer statistically significant when including the control variables.

5.3 Continent Specific Effects

As previously mentioned, continent specific effects is analyzed to add to the understanding of the subject since it could potentially be an element that has an influential effect on the idiosyncratic volatility. It is analyzed through regressions and the specification procedure is conducted in the same way as for the earlier regressions, resulting in the following regression models:

$$ln(IV)_{i,t} = \alpha + \beta_{con} Coninent_{i,t} + \sum_{x=1}^{10} \beta_x CV_{x,i,t} + \lambda_t + \mu_i + v_{it}, \qquad (13)$$

$$ln(IV)_{i,t} = \alpha + \beta_{Dev} Developed_{i,t} + \sum_{x=1}^{10} \beta_x CV_{x,i,t} + \lambda_t + \mu_i + v_{it}, \qquad (14)$$

where: α is the panel regression intercept; the betas are coefficients to the corresponding explanatory variables; $CV_{x,i,t}$ represents control variable x (of the in total 10 control variables), for firm i at time t; λ_t is a time-varying intercept that captures all of the variables that affect $ln(IV)_{i,t}$ and that vary over time but are constant cross-sectionally; μ_i is a variable that is constant over time but varies cross-sectionally in order to capture firm-specific effect; and v_{it} is the remainder disturbance term. Note that the model, explained by Equation (13), is regressed once for each of the five prespecified continents separately. The Developed variable in Equation 14 refers to the sales percentage in developed continents as explained in section 4.8.

The regressions, shown below in Table 9, indicate that *South & Central America* is the only continent, which has a statistically significant effect on idiosyncratic volatility. As can be seen from the continent coefficients, *Europe* is the only continent that has a negative relationship with the idiosyncratic volatility. When combining *Europe* and *NAFTA* in the same measure to form the *Developed* variable, the coefficient is negative. It is even more negative than when the continents were considered separately even though the coefficient is still not significantly different from zero.

Table 9: Continent Specific Effects Regressions

*Table 9 shows each of the prespecified continent specific effects regression models, using fixed effects in both dimensions and relevant control variables. T-statistics with the notation of a *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.*

	(1) Africa & Middle East	(2) Asia & Oceanic	(3) Europe	(4) NAFTA	(5) South & Central America	(6) Developed
α	-7,8460	-7,8210	-7,6922	-7,8016	-7,8761	-7,6125
T-Stat	*-20,14****	*-19,21****	*-19,23****	*-20,43****	*-20,65****	*-17,35****
Africa & Middle East	0,0083					
T-Stat	*0,60*					
Asia & Oceanic		0,0011				
T-Stat		*0,24*				
Europe			-0,0020			
T-Stat			*-0,62*			
NAFTA				0,0013		
T-Stat				*0,35*		
South & Central America					0,0339	
T-Stat					*2,05***	
Developed						-0,0030
T-Stat						*-0,78*
Age	0,0608	0,0517	0,0603	0,0453	0,0682	0,0690
T-Stat	*0,54*	*0,46*	*0,53*	*0,41*	*0,61*	*0,61*
Analysts	-0,0075	-0,0076	-0,0069	-0,0070	-0,0083	-0,0078
T-Stat	*-0,97*	*-0,97*	*-0,89*	*-0,90*	*-1,06*	*-1,00*
Beta	-0,0040	-0,0044	-0,0055	-0,0055	-0,0085	-0,0040
T-Stat	*-0,10*	*-0,11*	*-0,14*	*-0,14*	*-0,22*	*-0,10*
Div	-0,3428	-0,3311	-0,3253	-0,3440	-0,3498	-0,3124
T-Stat	*-2,16***	*-1,99***	*-2,02***	*-2,16***	*-2,24***	*-1,91**
Lev	0,1468	0,1322	0,1321	0,1270	0,1332	0,1410
T-Stat	*0,58*	*0,52*	*0,53*	*0,51*	*0,53*	*0,56*
M/B	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
T-Stat	*0,22*	*0,01*	*0,00*	*0,12*	*0,12*	*-0,09*
Price	-0,0013	-0,0013	-0,0013	-0,0013	-0,0013	-0,0013
T-Stat	*-2,62****	*-2,62****	*-2,64****	*-2,59****	*-2,72****	*-2,67****
ROA	0,4033	0,3936	0,4073	0,3994	0,4233	0,4006
T-Stat	*1,45*	*1,42*	*1,44*	*1,43*	*1,51*	*1,44*
Size	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
T-Stat	*0,73*	*0,74*	*0,72*	*0,76*	*0,85*	*0,71*
TO	-5,1765	-5,2248	-5,3804	-5,2994	-4,9848	-5,2448
T-Stat	*-1,36*	*-1,37*	*-1,42*	*-1,39*	*-1,31*	*-1,38*
R^2	0,8178	0,8177	0,8180	0,8178	0,8188	0,8179
Adj. R^2	0,7902	0,7900	0,7904	0,7901	0,7913	0,7903

6. Discussion

The purpose of the study is set out to assess how a firm's customer-base concentration impacts its stock return idiosyncratic volatility. More specifically, the research investigates the way that the distribution of sales, both geographically and operationally, influences the firm specific risk, i.e. the riskiness of a company's stock that is unrelated to the risk of the market.

As there is no universal standard for how customer-base concentration should be measured, it tends to be measured in a number of different ways. This research uses two different measures to represent the customer-base concentration. The first measure is based on the distribution of company sales among geographical segments and the second measure is based on the distribution of company sales among operating segments. In contrast to the measure based on geographical segment sales, the measure based on operating segment sales has been used in previous research including Dennis and Strickland (2009). Both measures are based on the theory that diversification across different markets influences firm-specific risk. Though the operational segments are defined according to how the companies have reported themselves, the geographical segments have been recategorized in terms of continents in order to obtain a measure that is more comparable across firms. The recategorization used for the geographical method creates more objectivity since the measure is no longer subject to how the firms define their geographical segments. Even though the recategorization creates a measure that is more objective and comparable across firms, it comes with a tradeoff. Due to the differences in how the geographical segments are defined by the companies, the new categorizes are rather wide (continent-based) leading to a certain loss of information as country specific sales that belong to the same continent are lumped together. In other words, diversification within the same continent is not accounted for. It can be argued that the objectivity and comparability gained from the recategorization outweighs the loss of information limitation that it is causing since the different firms report their geographical segments with rather varied level of specification. Therefore, the recategorization is applied to the geographical measure in this paper.

Based on the previous literature and the underlying theory it was hypothesized that the investigation would produce results indicating that there is a positive relationship between customer-base concentration and idiosyncratic volatility. The main findings of the research are that there is a relationship between the customer-base concentration and the idiosyncratic volatility that seems to be positive and significant at a first glance, but when including relevant control variables a relationship is no longer apparent. The univariate t-tests as well as the regressions excluding the control variables indicate that there is a positive and significant relationship between customer-base concentration and the idiosyncratic volatility for both measures of CBC. Despite this evidence of a potentially significant relationship, it is not possible to state with certainty that a higher CBC contributes to a higher idiosyncratic volatility since no relationship is observable after the inclusion of the control variables. Overall, the results indicate that the operating measure of CBC seems to have the strongest relationship with idiosyncratic volatility out of the two measures as it records higher coefficients in the regressions and more significant statistics in t-tests. The geographical measure of CBC even turns negative when the control variables are included in the regressions. Even though the operating CBC seemingly has a positive (and significant when the control variables are yet to be incorporated) relationship with the idiosyncratic volatility according to the results, both measures' relationships with idiosyncratic volatility are still insignificant with the inclusion of the control variables. Therefore there is not enough statistical evidence to draw any definite conclusions about the relationships between the variables.

The first tests conducted in the paper are two different types of univariate t-tests. One of these compares the average idiosyncratic volatility for firms with below median CBC to those above median CBC (see Table 4 and Table 5). The results show that there is a significant difference in average idiosyncratic volatility between the two halves. This would suggest that a higher CBC score is correlated with a higher idiosyncratic volatility, which is consistent with the hypothesis as well as the previous research made by Mihov and Naranjo (2017) who conducted a similar univariate t-test using another measure of customer-base concentration. Though it gives a quick indication of if there are any signs of a relationship between the variables, it is a rather simple test with a number of limitations. One of the limitations is that it does not provide any insight to the dynamics of the relationship but rather only indicates if there is a relationship and whether or not it is positive or negative. Another limitation is that splitting up the firms into two halves based on their CBC does not necessarily ensure any major differences between the two groups as most firms could have a CBC-score that lies close to the median value. Additionally, the results could be misleading as it does not control for the fact that there might be other variables that could be causing the difference in idiosyncratic volatility between the groups.

The other univariate t-test divides the firms into four CBC quartiles and compares the average idiosyncratic volatility between the lowest and the highest CBC quartile. This test has the same limitations as the previous univariate t-test, but it solves the issue of comparing two groups with firms that have a CBC-score close to the median; it ensure that there is a larger difference in CBC between the two groups compared, which according to the theory should produce more significant results. Using this test, similar results are obtained for the operating CBC measure but not for the geographical CBC measure (See Table 6 and Table 7). The results from the test using the operating CBC reinforces the idea that a higher concentration leads to a higher IV as it progressively increases each quartile for most of the separate years. In line with the results of Mihov and Naranjo (2017), more significant results are shown when comparing the first and last quartile than when comparing the two halves. On the other hand, the results from the test using the geographical CBC are not as clear cut. The idiosyncratic volatility does increase progressively for the first three quartiles but then experience a notable decrease for the fourth quartile where the firms with a CBC-score of 1 are included. There are a few different ideas that could potentially explain these results: the actual relationship is in line with the hypothesis and it is merely a coincidence specific to the sample that the firms in the fourth quartile have slightly lower IV than the previous quartile; the actual relationship is in line with the hypothesis but the geographical CBC measure is not specific enough relating to the loss of information issue discussed earlier; the hypothesis does not apply for geographical diversification and thus no significant results are obtained; or there might be a other factors that are directly related to the firms with a CBC score of 1 that is causing the decrease in the fourth quartile. Since results are consistently in line with the hypothesis up until the fourth quartile and there is a large difference in idiosyncratic volatility between the first and the third quartile, the idea that there is something peculiar about the fourth quartile is definitely a possibility based on the results from the univariate t-tests.

Following the univariate t-tests, three different regression models are run both with and without the control variables (see Table 8). Through running the regressions, the complete dynamics of the relationship can be better understood and by introducing relevant control variables the causal factors of idiosyncratic volatility can be thoroughly examined. When analyzing the results from the regressions where the control variables are excluded it can be observed that both the geographical and operating CBC measures are significantly and positively related to the idiosyncratic volatility when regressed independently. This supports the idea that both the geographical and the operating customer-base concentration increases the idiosyncratic volatility meaning that distributing sales both geographically and among operating segments has a diversifying effect that lowers the firm specific risk. When both measures are regressed simultaneously, the positive relationships still appears to remain; however, the operational CBC takes away a substantial portion of the effect that the geographical CBC has on idiosyncratic volatility as the coefficient of the geographical CBC is nearly cut in half and is no longer significantly different from zero. Though some of the effect is also lost for the operating CBC measure, it remains positive and statistically significant. Similarly to the univariate t-tests, these results once again indicate that the relationship between the CBC and idiosyncratic volatility is stronger for the operating measure. Compared to the univariate t-tests, these regressions explain more about the dynamics between the two measures of customer-base concentration and idiosyncratic volatility, even though it still does not control for other firm specific factors that might be causing the results.

When analyzing the results from the regressions where the control variables are included, a significant relationship between customer-base concentration and idiosyncratic volatility is no longer observable (see Table 8). The control variables that are included in the regressions are variables that have been shown to have a significant effect on idiosyncratic volatility in previous literature. Out of the 10 control variables applied, some appears to be consistent with the theory in that they have the same sign as the previous studies suggest. All of those that appear to be inconsistent with previous studies have insignificant coefficients and it is therefore difficult to draw any conclusions about whether or not they go against the findings of previous studies. Two of the variables have coefficients that are significantly different from zero. These two variables are *Div* and *Price* and they appear to be in accordance with the findings of previous studies. *Div* has a significantly negative coefficient meaning that dividend-paying firms have less volatile returns, which is also found by Pástor and Veronesi (2003). *Price*, which controls for the impact of microstructure noise in low priced stocks, also has a significant negative coefficient which lies in accordance with the findings of Gaspar and Massa (2006).

As mentioned, the CBC variables go from being significant to insignificant in trying to explain the idiosyncratic volatility when all control variables are accounted for in the regressions. The effect that is found in the previous results is thus not attributable to the customer-base concentration but instead to other firm specific factors explained by the control variables. It is also noticeable that the operating CBC still has a positive coefficient while the coefficient of the geographical CBC turns negative in both regressions. This is once again consistent with the idea that the operational measure has a larger influence on idiosyncratic volatility than the geographical measure. Dennis and Strickland (2009), who use the same measure for customer base concentration as the operating CBC used in this paper, obtained significant results for their customer base measure in explaining the development of aggregate idiosyncratic volatility over time. Because of this it comes no surprise that the operational CBC measure repeatedly appears to have a positive influence on idiosyncratic volatility. In contrast, the geographical CBC measure is an untested measure completely motivated by a theoretical framework and yet to be applied in other literature and it is therefore not completely unremarkable that a significant effect on idiosyncratic volatility is not found. The geographical CBC measure could potentially also be limited by the phenomenon explained in the theory presented by Fillat and Garetto (2015) who argue that companies with a multinational customer-base have a larger cash-flow risk exposure since firms are hesitant to withdraw from markets abroad due to geographical market penetration sunk costs. Shown by Koller, Goedhart and Wessels (2015), this cash-flow risk exposure is directly related to the idiosyncratic volatility (see Section 2.4 for full explanation). This concept could therefore potentially explain the negative coefficient of the geographical CBC measure and the overall insignificant effect of the variable on the idiosyncratic volatility.

Lastly, the geographical customer-base concentration is further investigated through a regression that aims to demonstrate if there are any continent specific effects on idiosyncratic volatility (see Table 9). Even though it is not statistically significant, *Europe* is the only continent variable that has a negative coefficient suggesting that having sales in this region potentially lowers idiosyncratic volatility. This could explain the results observed in the univariate t-test where the firms with a geographical CBC score of 1 experience lower idiosyncratic volatility, which do not follow the pattern found between the other quartiles (see Table 6). In the sample, all the firms with a CBC score of 1 have sales completely centered in Europe and if there is a continent specific effect for Europe that lowers idiosyncratic volatility then this helps to explain why the idiosyncratic volatility drops for the firms with a CBC score of 1. In addition, this could potentially also explain why the geographical CBC coefficients are negative in the regressions when the control variables are included. Since only Swedish companies are used in the investigation, the firms with high customer-base concentration are disproportionally based in Europe meaning that, in most of the cases, a higher geographical CBC score also implies a higher concentration in Europe. Thus, if there is a continent specific effect for Europe that lowers idiosyncratic volatility then this possibly demonstrates why the results in the regressions for the geographical CBC is inconsistent with the underlying theory when including the control variables (see Table 8). When combining *Europe* and *NAFTA* in the same measure to form the *Developed* variable, the coefficient becomes even more negative though it is still not significantly different from zero. This is consistent with the previous study by Angelidis (2010), which argues that the behavior of idiosyncratic volatility differs between developed and emerging markets. Although the signs of the coefficients are consistent with the idea that the idiosyncratic volatility is lower in developed markets as opposed to emerging markets, the coefficients are not significantly different from zero meaning that the results do not provide enough statistical evidence to make this conclusion with certainty.

Moreover, a statistically significant relationship between customer-base concentration and idiosyncratic is not found when including the control variables and the results are thus not able to confirm the underlying theory and findings of the previous literature in the field. This could perhaps be explained by the limitations of the study, one of which is the sample size. A lot of the variables are insignificant when regressed simultaneously and though this could be because that the actual variables have no relationship to the idiosyncratic volatility, it might as well be that the current sample size is not extensive enough to produce standard errors that lead to precise statistical inferences. Another limitation is the information loss when constructing the geographical CBC measure. The imprecision of the geographical CBC measure, caused by making the measure objective and comparable across firms, is another factor that could explain the insignificance of the geographical CBC when regressed with the control variables; it could perhaps also explain why it consistently has a lower effect on idiosyncratic volatility compared to the operating CBC measure. Furthermore, as the sample only includes Swedish companies the generalizability of the results might be limited to firms in this region and not applicable to firms worldwide. Having only Swedish companies in the sample also creates a potential continent specific effects issue since the firms with a high concentration are disproportionally based in Europe meaning that a higher geographical CBC score also implies a higher concentration in Europe.

To counteract the limitations of the study, a number of extensions for further research within the subject are suggested. Firstly, increasing the sample size and incorporating firms from all over the world would make the research both more generalizable and help reduce the standard errors in order to obtain more significant results. Secondly, though no significant evidence for continent specific effects is found in this paper, this is something that could be investigated further and perhaps be accounted for. Using a firm sample that is evenly distributed among different continents would reduce this potential effect. Finally, one could also use a more specific geographical CBC measure by recategorizing the sales segments in terms of countries instead of continents. This would require a sample containing firms that all report country specific sales and this suggestion thus reflects back to the extension of using a more comprehensive sample.

7. Conclusion

The study sets out the objective to investigate how a firm's customer-base diversification affects its stock return idiosyncratic volatility. The research paper mainly focuses on the way that the distribution of sales impacts the firm specific volatility, which is unrelated to the risk of the market. It does so through the use of both a geographical and an operating customer-base concentration measure. Previous literature and the underlying theory mainly propose that there is a positive relationship between customer-base concentration and idiosyncratic volatility. The findings of this study appear to be in line with that hypothesis at a first glance; however, the results lack significance when controlling for relevant firm specific factors. Both the univariate t-test and the regressions that exclude the control variables find a positive and significant effect between customer-base concentration and idiosyncratic volatility. This suggests that sales distribution in terms of geographical and operating segments has a diversification effect on idiosyncratic volatility. The significance of the effect is however absorbed by the control variables when introduced to the models. This proposes either that other firm-specific factors are the cause of the increase in idiosyncratic volatility or that the sample size is too limited to provide significant evidence supporting a relationship between customer-base concentrations and idiosyncratic volatility. Furthermore, the results demonstrate that the operating CBC measure repeatedly has a stronger effect on idiosyncratic volatility than the geographical CBC measure. Though this might propose that operating diversification is more effective than geographical diversification, the relatively weak effect found of the geographical CBC measure is potentially just a result of continent specific effects or the loss of information when recategorizing.

8. Appendix

Table 1: List of Firms

Table 1 presents the 60 firms used in the investigation.

AAK	Haldex	Orexo
Alfa Laval	Hennes & Mauritz	Ratos
Assa Abloy	Holmen	Rottneros
AstraZeneca	ICA	SAAB
Atlas Copco	Indutrade	Sandvik
Atrium Ljungberg	Intrum Justitia	SCA Svenska Cellulosa
Avanza	JM	Securitas
Axfood	Kungsleden	Skanska
Beijer Alma	Lundin Petrolium	SKF
Beijer Ref	Meda	SSAB
Betsson	Medivir	Starbreeze
Bilia	Mekonomen	Svenska Handelsbanken
BioGaia	Modern Times Group	Sweco
Boliden	Mycronic	Tele2
Castellum	NCC	Telefonaktiebolaget Ericsson
Clas Ohlson	Net Insight	Telia
Eniro	New Wave	Trelleborg
Fabege	Nobia	Wallenstam
Fastighets Balder	Nolato	Wihlborgs
Fingerprint	Nordea	Volvo

Figure 1: IV against Firm and Year

Figure 1 shows the IV for each firm and year.

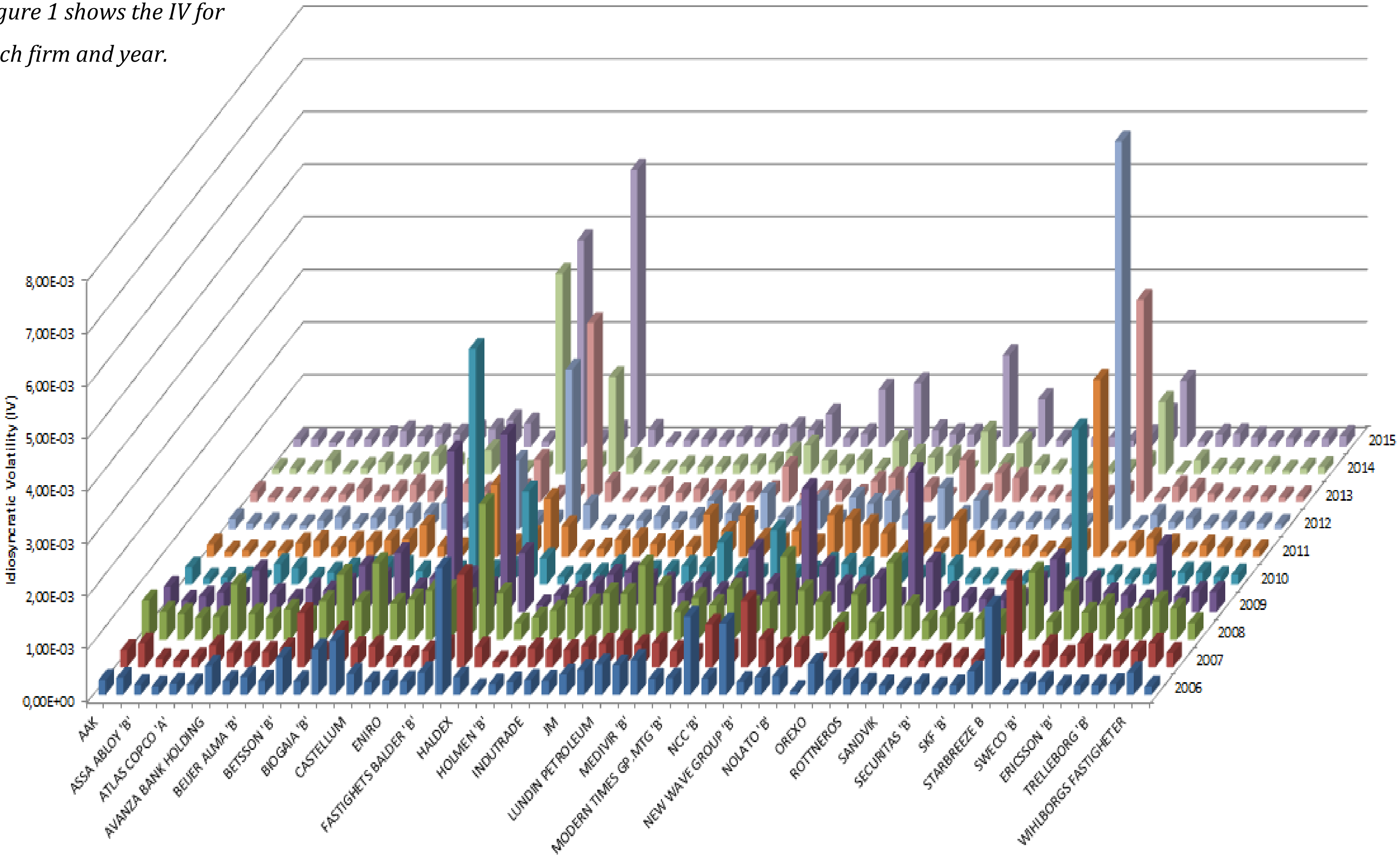

Figure 2: Mean & Median IV

Figure 2 shows the mean and median of IV over time.

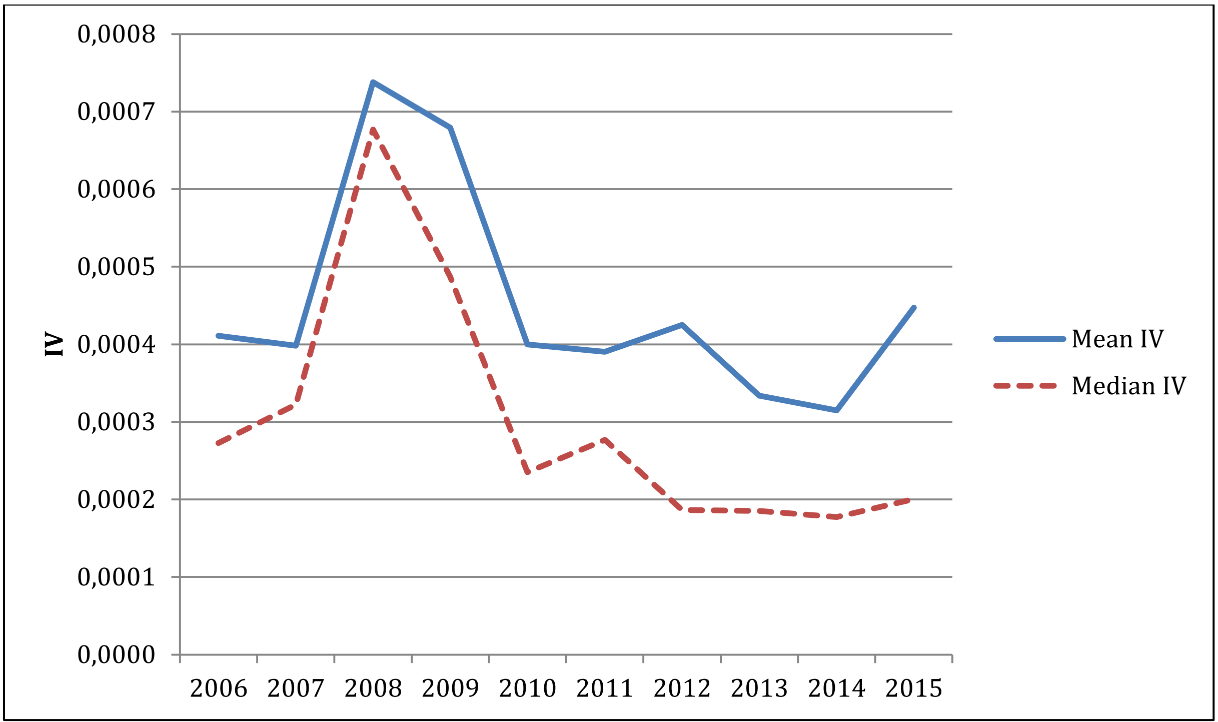

Figure 3: Mean & Median ln(IV)

Figure 3 shows the mean and median of ln(IV) over time.

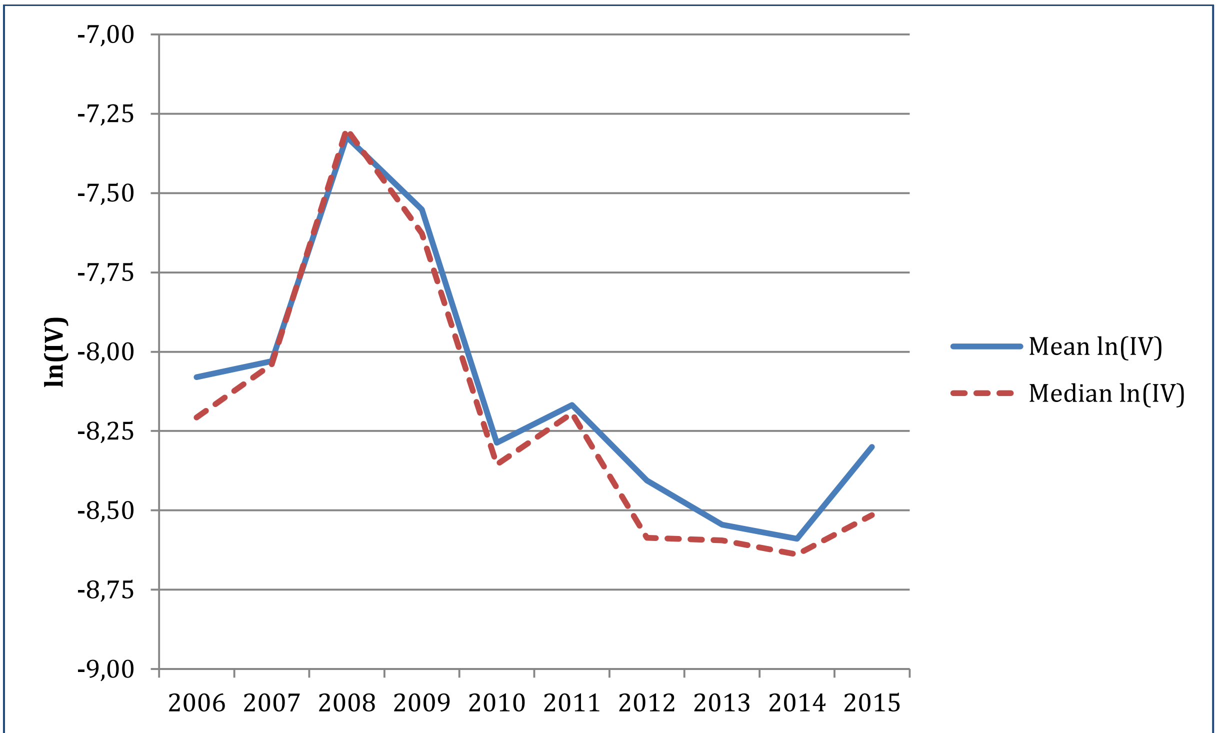

Table 2: Correlation Matrix

Table 2 demonstrates the correlation matrix between ln(IV), CBC_{Op}, CBC_{Geo}, *and all control variables.*

	$\ln(IV)$	CBC_{Op}	CBC_{Geo}	Age	Analysts	Beta	Div	Lev	M/B	OMXS30	Price	ROA	Size
CBC_{Op}	0.17	-											
P-Value	*0.00*												
CBC_{Geo}	0.13	0.46	-										
P-value	*0.00*	*0.00*											
Age	-0.40	-0.23	-0.29	-									
P-Value	*0.00*	*0.00*	*0.00*										
Analysts	-0.48	-0.24	-0.31	0.49	-								
P-Value	*0.00*	*0.00*	*0.00*	*0.00*									
Beta	-0.11	-0.12	-0.16	0.30	0.48	-							
P-Value	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*								
Div	-0.54	-0.28	-0.04	0.29	0.32	0.11	-						
P-Value	*0.00*	*0.00*	*0.34*	*0.00*	*0.00*	*0.01*							
Lev	-0.17	-0.01	0.12	0.01	0.12	0.10	0.18	-					
P-Value	*0.00*	*0.86*	*0.00*	*0.89*	*0.00*	*0.01*	*0.00*						
M/B	0.01	0.03	-0.01	-0.01	0.05	0.02	-0.10	0.13	-				
P-Value	*0.78*	*0.47*	*0.85*	*0.74*	*0.19*	*0.58*	*0.01*	*0.00*					
OMXS30	-0.41	-0.25	-0.40	0.44	0.83	0.42	0.22	0.03	0.06	-			
P-Value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.46*	*0.16*				
Price	-0.40	-0.14	-0.08	0.18	0.15	-0.02	0.31	-0.06	0.01	0.21	-		
P-Value	*0.00*	*0.00*	*0.05*	*0.00*	*0.00*	*0.63*	*0.00*	*0.17*	*0.79*	*0.00*			
ROA	-0.22	-0.04	-0.04	0.07	0.08	-0.07	0.33	-0.17	-0.07	0.06	0.23	-	
P-Value	*0.00*	*0.38*	*0.35*	*0.07*	*0.06*	*0.09*	*0.00*	*0.00*	*0.07*	*0.13*	*0.00*		
Size	-0.41	-0.20	-0.24	0.43	0.61	0.15	0.20	-0.04	0.00	0.61	0.37	0.15	-
P-Value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.32*	*0.93*	*0.00*	*0.00*	*0.00*	
TO	-0.05	-0.13	0.00	0.06	0.26	0.26	0.14	0.06	0.02	0.35	0.48	0.11	0.05
P-Value	*0.19*	*0.00*	*0.99*	*0.12*	*0.00*	*0.00*	*0.00*	*0.16*	*0.59*	*0.00*	*0.00*	*0.01*	*0.27*

Table 3: Unit Root Test

Table 3 demonstrates the augmented Dickey Fuller test statistics and corresponding p-values for ln(IV), CBC_{Op}*,* CBC_{Geo}*, and all control variables except for the dummy variables Div and OMXS30.*

	T-Stat	P-Value
ln(IV)	240.09	*0.0000*
CBC(Op)	180.90	*0.0000*
CBC(Geo)	132.14	*0.0002*
Age	958.18	*0.0000*
Analyst	165.32	*0.0008*
Beta	255.74	*0.0000*
Lev	177.00	*0.0001*
M/B	291.76	*0.0000*
Price	178.13	*0.0005*
ROA	261.26	*0.0000*
Size	173.28	*0.0011*
TO	212.42	*0.0000*

Table 4: Heterogeneity Test in the Cross-sectional Dimension

Table 4 shows the F-statistic and Chi-Square statistic with corresponding p-values resulting from a fixed effects redundancy test when using fixed effects in the cross-sectional dimension. The test is performed on all three regression models, where $Regression_1$ *contains all variables except for* $CBC_{Op,t}$*,* $Regression_2$ *contains all variables except for* $CBC_{Geo,t}$*, and* $Regression_3$ *contains all variables.*

	F-Statistic	P-Value	Chi-Square Statistic	P-Value
$Regression_1$	5.11	*0.0000*	270.69	*0.0000*
$Regression_2$	5.12	*0.0000*	270.95	*0.0000*
$Regression_3$	5.10	*0.0000*	270.57	*0.0000*

Table 5: Heterogeneity Test in the Time Dimension

Table 5 shows the F-statistic and Chi-Square statistic with corresponding p-values resulting from a fixed effects redundancy test when using fixed effects in the period dimension. The test is performed on all three regression models.

	F-Statistic	P-Value	Chi-Square Statistic	P-Value
$Regression_1$	25.16	*0.0000*	198.04	*0.0000*
$Regression_2$	25.25	*0.0000*	198.63	*0.0000*
$Regression_3$	25.19	*0.0000*	198.53	*0.0000*

Table 6: Combined Heterogeneity Test in both Dimensions

Table 6 shows the F-statistic and Chi-Square Statistic with corresponding p-values resulting from a fixed effects redundancy test when having fixed effects in both dimensions simultaneously. The test is performed on all three regression models.

	Cross-section			
	F-Statistic	P-Value	Chi-Square Statistic	P-Value
$Regression_1$	7.95	*0.0000*	386.22	*0.0000*
$Regression_2$	7.93	*0.0000*	385.15	*0.0000*
$Regression_3$	7.94	*0.0000*	385.54	*0.0000*
	Period			
	F-Statistic	P-Value	Chi-Square Statistic	P-Value
$Regression_1$	39.68	*0.0000*	314.18	*0.0000*
$Regression_2$	39.49	*0.0000*	312.50	*0.0000*
$Regression_3$	39.61	*0.0000*	313.23	*0.0000*

Table 7: Endogeneity Test for Random Effects Models

Table 7 shows the Chi-Square statistics resulting from a Hausman test on the regressions using a random effects specification. The test is performed on all three regression models to identify potential endogeneity problems.

	Chi-Square Statistic	P-Value
$Regression_1$	53.02	*0.0000*
$Regression_2$	52.48	*0.0000*
$Regression_3$	54.46	*0.0000*

9. References

Angelidis, T. (2010). Idiosyncratic Risk in Emerging Markets. *The Financial Review*, 45(4), pp.1053-1078.

Bartram, S., Brown, G. and Waller, W. (2012). How Important is Financial Risk. *Journal of Financial and Quantitative Analysis*, 50(4), pp.801-824.

Bates, T., Stultz, R. and Kahle, K. (2009). Why do U.S. firms hold so much more cash than they used to?. *The Journal of Finance*, 64(5), pp.1925-2021.

Black, F. (1976). *Studies of stock price volatility changes*. Proceedings of the 1976 Meetings of the Business and Economic Statistics Section. American Statistical Association, pp. 177–181

Black, F. and Scholes, M. (1973). The Pricing of Option and Corporate Liabilities. *The Journal of Political Economy*, 81(3), pp.637-654.

Brooks, C. (2014). *Introductory Econometrics for Finance*. 3rd ed. Cambride: Cambridge University Press, pp.217-220.

Chay, J and Suh, J. (2009). Payout Policy And Cash-Flow Uncertainty. *Journal of Financial Economics*, 93(1), pp.88-107.

Dennis, P. and Strickland, D. (2009). *The Determinants of Idiosyncratic Volatility*. Charlottesville: University of Virginia.

Ernst & Young (2009). *IFRS 8 Operating Segments - Implementation Guidance*. Global: Ernst & Young.

Fillat, J. and Garetto, S. (2015). Risk, Returns, and Mulitnational Production. *The Quarterly Journal of Economics*, 130(4), pp.2027-2073.

Gaspar, J. and Massa, M. (2006). Idiosyncratic Volatility and Product Market Competition. *The Journal of Business*, 79(6), pp.3125-3152.

Goetzmann, W. and Kumar, A. (2008). Equity Portfolio Diversification. *Review of Finance*, 12(3), pp.433-463.

Goyal, A. and Santa-Clara, P. (2003). Idiosyncratic Risk Matters. *The Journal of Finance*, 58(3), pp.975-1007.

Hoberg, G. and Prabhala, N. (2009). Disappearing Dividends, Catering, and Risk. *The Review of Financial Studies*, 22(1), pp.79-116.

Irvine, P. and Pontiff, J. (2009). Idiosyncratic Return Volatility, Cash Flows, and Product Market Competition. *The Review of Financial Studies*, 22(3), pp.1149-1177.

Koller, T., Goedhart, M. and Wessels, D. (2015). *Valuation: Measuring and Managing the Value of Companies*. 6th ed. New Jersey: John Wiley & Sons, pp.17-33.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), pp.13-37.

Malkiel, B. and Xu, Y. (2001). *Idiosyncratic Risk and Security Returns*. Working Paper. Dallas: University of Texas at Dallas.

Malkiel, B. and Xu, Y. (2003). Investigating the Behavior of Idiosyncratic Volatility. *The Journal of Business*, 76(4), pp.613-644.

Markarian, G. and Gill-de-Albornoz, B. (2012). Income Smoothing and Idiosyncratic Volatility. Madrid: SSRN.

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

Mihov, A. and Naranjo, A. (2017). Customer-base concentration and transmission of idiosyncratic volatility along the vertical chain. *Journal of Empirical Finance*, 40, pp.73-100.

Morck, R., Yeung, B. and Yu, W. (2000). Why Do Emerging Markets Have Synchronous Stock Price Movements?. *The Journal of Financial Economics*, 58(1-2), pp.215-260.

Pastor, L. and Veronesi, P. (2003). Stock Valuation and Learning about Profitability. *The Journal of Finance*, 58(5), pp.1749-1789.

Patatoukas, P. (2012). Customer-Base Concentration: Implications for Firm Performance and Capital Markets. *The Accounting Review*, 87(2), pp.363-392.

Polkovnichenko, V. (2005). Household Portfolio Diversification: A Case for Rank-Dependent Preferences. *The Review of Financial Studies*, 18(4), pp.1467-1502.

Redovisningsrådet (2002). *Rapportering för segment – rörelsegrenar och geografiska områden*. Redovisningrådets Rekommendationer. [online] Bokföringsnämnded, pp.3-4. Available at: http://www.bfn.se/upload/sv/redovisningsregler/redovisningsradets-rekommendationer/RR/RR25.pdf [Accessed 18 Apr. 2017].

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), pp.425-442.

United Nations Statistics Division (2017). *Methodology: Standard country or area codes for statistical use (M49)*. [online] Available at: https://unstats.un.org/unsd/methodology/m49/ [Accessed 18 Apr. 2017].

Vuolteenaho, T. (2002). What drives firm-level stock returns?. *The Journal of Finance*, 57(1), pp.233-264.

Wurgler, J. and Zhuravskaya, E. (2002). Does Arbitrage Flatten Demand Curves for Stocks. *Journal of Business*, 75(4), pp.583-608