The returns, risk and liquidity relationship in high frequency trading: Evidence from the Oslo stock market

Minh Thi Hong Dinh[1]

The School of Business and Economics at UIT – The Arctic University of Norway

A R T I C L E   I N F O

A B S T R A C T

The main purpose of this research is to investigate the relationship between returns, risk, and liquidity in high frequency trading. Panel analysis for single stocks is employed to investigate this relationship. The empirical results imply that in high frequency trading idiosyncratic risk plays a more pronounced role than systematic risk in asset pricing. First, idiosyncratic risk and liquidity have a highly significant impact on returns. Second, no evidence has been found for a significant relationship between systematic risk and returns. Finally, liquidity has a higher significant effect on idiosyncratic risk than systematic risk. The empirical results of the paper contribute to the previous literature in the high frequency context. Some previous literature suggests that idiosyncratic risk has a matter on low frequency trading, but has not yet investigated its effects on high frequency trading.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The relationship between risk and returns is a major topic, discussed by many researchers. This relationship is recognized in the CAPM theory developed by Sharpe (1964), Lintner (1965), and Mossin (1966). An underpinning notion of this theory is that a diversified portfolio of stocks is less risky than any of its components (Mullins, 1982). Aggregate risk includes specific and systematic risk; the effects of specific risk are reduced as more securities are added to a portfolio. An investor holding a properly diversified portfolio is therefore not compensated for specific risk, but only for the systematic. However, Merton (1973) and Boehme et al. (2009) argue that investors cannot hold a properly diversified portfolio because of incomplete stock information and the existence of many different costs in the financial market. Investors are willing to invest in stocks that they are familiar with. Idiosyncratic risk should therefore be priced in predicting stock returns when investors hold a diversified portfolio, but not one which is a properly diversified, and the relationship between them could be positive. Mullins (1982) also suggests that corporate securities move together to some extent in the financial market, so the complete

elimination of specific risk from a simple portfolio is impossible. Hence, idiosyncratic risk should be priced in cross-sectional stock returns.

The goal of this research is to provide an examination of how aggregate volatility risk is priced in the cross-section of stock returns. The differences between this research and previous work are that the past literature does not investigate idiosyncratic risk at the intraday level and does not employ the panel data of individual stocks for analysis. The highest frequency level of data used to estimate idiosyncratic risk is found at the daily level in the research of Ang et al. (2006), while idiosyncratic risk in this research is estimated by using intraday data. This research does not examine the relationship between risk and returns by sorting stocks into portfolios as the previous literature does, but explores the panel analysis of individual stocks. Although there are some differences between this study and previous ones, the results of it are consistent with previous ones. For example, finding idiosyncratic risk in a negative relation with returns is consistent with findings of the studies by Ang et al. (2006, 2009).

This research contributes to the previous literature in the following ways. It first employs variables at a higher frequency than past research does. It uses the daily variables estimated from intraday ones which are at one minute intervals. Several prior empirical studies indicate that idiosyncratic risk matters in low frequency trading, but its effect on high frequency trading has not yet to be investigated. Second, the use of panel data analysis rather than portfolio analysis is another contribution of the paper. Third, this research answers the question of the extent to which liquidity contributes to pricing expected returns, by addressing the risk and liquidity relationship. Moreover, it supports the findings of Ang et al. (2006, 2009) that idiosyncratic risk and returns are in an inverse relationship, and of other past studies which found that systematic risk does not affect returns. Finally, finding that liquidity variables affect idiosyncratic risk implies that when this risk is considered in pricing stock returns, liquidity factors should be also considered.

This research is important for two reasons. First, high frequency trading has become popular nowadays, but the existing literature on the risk and returns relationship is largely focused on low frequency trading. It could be of great value to extend the existing empirical work to high frequency trading. Finding the relationship between these factors in high frequency trading is important for all market participants as it would help them to manage their trading portfolios more efficiently. Finally, the existing empirical work uses traditional portfolio analysis. It would be useful to examine the relationship by employing the method of panel analysis for single stocks. The reason for this is that individual stocks interact together to some extent cross-sectionally and over time. Adding individual stocks to a portfolio for analysis would limit their interaction.

The results of this research indicate that idiosyncratic risk plays a more important role than beta in high frequency trading when pricing expected returns. This can be summarized as follows. First, there is a negative relation between returns and idiosyncratic risk, similar to that in some prior literature. Second, the relationship between returns and liquidity is positive. Third, a flat relation between beta and returns is suggested, which is consistent with the previous empirical studies. Finally, liquidity variables affect idiosyncratic risk more than beta. The positive or negative effects depend on how the liquidity variables are defined.

The remainder of this paper proceeds as follows: in section two, previous related empirical literature is reviewed. Section three describes the data, variables, panel analysis and panel data. In section four, two models, the empirical results are presented, and the some discussions are in the next section. The conclusions are addressed in the final section.

2. Literature review

There is a vast body of literature that has examined the relationship between risk and stock returns. Much of it, however, employs low frequency data and documents different results. Many researchers suggest that the risk-return relationship is an inverse one, while some demonstrate that these factors have a positive relationship. However, Bali and Cakici (2008), and Berggrun et al. (2016) find no significant relationship between them.

A flat relationship between returns and idiosyncratic risk is found in the research by Bali and Cakici (2008), and Berggrun et al. (2016). Generating two different monthly idiosyncratic risks, first from the daily returns of previous months and second from 25 to 60 monthly returns, Bali and Cakici (2008) model the relationship between idiosyncratic risk and cross-section returns and apply the CAPM and the three factor Fama–French models to examine the relationship. They state that there is no strongly significant relationship between idiosyncratic risk and returns for both types of idiosyncratic risk. They indicate, however, that some other factors affect returns, such as data frequency, weighting schemes for calculating portfolio returns, breakpoints for sorting stocks, and a screen for size, price and liquidity. Berggrun et al. (2016) investigate this relationship by using stocks in the MILA, but they could not find any relationship between idiosyncratic risk and returns. They suggest that idiosyncratic risk is not a price factor in this market.

Some literature demonstrates that idiosyncratic risk and returns have a positive relationship. Employing a different method to Merton (1987) and Ang et al. (2006), Fu (2009) applies the exponential GARCH model to investigate the relationship between idiosyncratic risk and expected returns. He finds a positive relation between these factors. Fu also mentions that the research by Merton (1987) and Ang et al. (2006) should not infer that there is a relationship between idiosyncratic risk and expected returns because idiosyncratic volatility is time-varying. Malkiel and Xu (2002) also address the positive relation between idiosyncratic risk and returns by applying the framework of the Fama–MacBeth, and the Fama–French models. Their empirical results conclude that idiosyncratic risk is very useful variable in explaining cross-sectional expected returns. Another result found by Malkiel and Xu (2002) is that liquidity variables and idiosyncratic risk have an insignificant relationship. Goyal and Santa-Clara (2003) found that idiosyncratic risk and market returns have a positive relationship,

but they were unable to find the relationship between market variance and market returns. The topic of the risk-return relationship relative to IPO returns is also found in research by Wagner (2004). He indicates that the relationship between conditional idiosyncratic risk and expected returns is positive during the first two years after initial listing on the financial market.

A negative relationship between idiosyncratic risk and returns is suggested by some previous empirical work. Contrary to Fu (2009) and Malkiel and Xu (2002), Ang et al. (2006, 2009) show a negative relation between expected returns and idiosyncratic risk. Generating a stock portfolio by sorting stock sensitivities to innovation in total risk, Ang et al. (2006) first examine the effects of aggregate risk on cross-sectional returns based on the multi-factor models of Merton (1973) and Ross (1976). Second, by estimating idiosyncratic risk from the Fama–French model, they model the relationship between returns and idiosyncratic risk. Their empirical results suggest that stocks generate low returns if they have high sensitivities to innovation in total risk, and stocks with high idiosyncratic risk have extremely low returns. Ang et al. (2009) investigate this relationship in the US and international regions. Estimating idiosyncratic risk from the Fama-French model at three different levels, the local, regional, and global, Ang et al. (2009) also find that stocks with high idiosyncratic risk have low returns. Moreover, they find significant co-movement of low return stocks around the world with the effects of idiosyncratic risk in the US. Cremers et al. (2015) examine both aggregate jump and volatility risk in pricing cross-sectional stock returns. They suggest that stocks with high sensitivity to jump and volatility risk generate low expected returns. Hatemi-J and Irandoust (2011) examined the causal relationship between returns and volatility and found that volatility causes negative returns, but returns cause volatility positively.

Research by Huang et al. (2009) found that the relationship between idiosyncratic risk and expected stock returns is negative if the estimate is based on daily returns, but it is positive if the idiosyncratic risk is estimated from monthly data. In addition, the characteristics and forecasting ability of idiosyncratic risk is investigated by Angelidis and Tessaromatis (2008). They use data from the UK market to investigate this topic while most past studies use data from the US. They examine three types of idiosyncratic risk: value-weighted risk, idiosyncratic risk measures based on large capitalization stocks, and risk based on small capitalization stocks. Following Goyal and Santa-Clara (2003), they measure idiosyncratic risk based on variance of stocks and it is defined as the variance of the idiosyncratic returns. Their empirical results suggest that the predictive power of idiosyncratic risk is robust and remains significant, even after they manage for possible persistence in the risk. Storesletten et al. (2007) study what affects idiosyncratic risk by adding the life cycle and capital accumulation into the relationship between asset-market risk premiums and idiosyncratic shocks. They show that these two ingredients mitigate the effects of idiosyncratic risk relative to the observed Sharpe ratio on US equity, and illustrate that idiosyncratic risk matters for asset pricing. The research on arbitrage and idiosyncratic risk by Pontiff (2006) illustrates that arbitrageurs face high costs from idiosyncratic risk in mispricing.

The literature review indicates that the effects of idiosyncratic risk on returns, which are currently generated from low frequency data, are different: negative, positive, with a flat relationship and with mixed relationships. The limitation of the previous empirical work is that it has not covered all aspects of trading in financial markets relative to the level of frequency. This means that past studies focus on low frequency. The present work allows us to test this relationship by using higher frequency data, which is the main contribution of this research.

3. Data, variables, panel analysis and panel data description

3.1. Data description

Intraday data from January 2003 to April 2010 obtained from the Oslo Stock Exchange (OSE) is used to generate daily variables. The sample comprises 116,583 daily observations for 150 companies. There is a substantial variation in both size and liquidity among the companies listed on the OSE. The largest companies have an average of 15,000 orders and approximately 3,000 trades per day. On the other hand, the smallest ones have low liquidity, with approximately 30 trades per day. All changes during a day for a particular company are recorded in an order book, which is re-constructed by algorithms by which the information on trading is added to, removed from, or modified on both the bid and ask sides of the order book. After that, snapshots are taken of all levels in the order book at 60s intervals.

3.2. Variables

Intraday data is used to generate daily variables in this study. The data consists of mid prices, which means the average between bid and ask prices, from the last change in the order book. The mid prices are collected every 60 seconds during an e-trading day. A market portfolio is generated each day, based on the market cap of each company that day.[2] Daily beta, idiosyncratic risk, and returns are then calculated from the mid prices. The CAPM model can be written:

$$R_{t,i,j} = \alpha_{i,t} + \beta_{t,i,j} R_{m,j,t} + \epsilon_{j,i,t} \tag{1}$$

---

[2] For some small illiquid companies with little or no trade, there was no available market cap, so it was set to 1e−11.

Table 1

Summary of the Panel Data

The panel data below covers the period from 2003 to 2010 and is filtered from the original data which comprised 116,583 daily observations, by screening out low liquidity stocks. Beta, returns, and residualVar are daily variables calculated from the intraday data based on the CAPM model. Idio (or idiosyncratic risk) is defined as logresidualVar. nTrades, Turnover, BindFreq are liquidity variables. The lidio, lbeta, and lagBindFreq are the idiosyncratic risk, systematic risk, and liquidity variables of the previous day respectively.

Statistic	N	Mean	St. dev.	Min	Max
Beta	18,287	0.831	0.554	−5.357	26.490
Returns	18,287	−0.00001	0.0001	−0.001	0.001
Residualvar	18,287	0.000003	0.00002	0.00000003	0.001
nTrades	18,287	2043.689	1411.727	836	19,847
Turnover (in thousands)	18,287	446,547	644,384	10,932	17,318,201
BindFreq	18,287	0.622	0.289	0.000	1.000
Idio	18,287	−13.580	1.173	−17.280	−6.688
Lidio	18,287	−13.628	1.229	−19.394	0.000
Lbeta	18,287	0.814	0.563	−2.804	26.490
LagBindFreq	18,287	0.619	0.291	0.000	1.000
Length	18,287	1484.009	198.144	1074	1791

where $R_{t,i,j}$ are asset log returns and $R_{m,j,t}$ are market log returns. The subscripts represent day t, stock i, and minute j. Standard deviations of the error terms $\epsilon_{j,i,t}$ represent intraday idiosyncratic risk, $\beta_{t,i,j}$ is the CAPM beta, and $\alpha_{t,i}$ represents idiosyncratic returns for a trading day for a given stock. Daily idiosyncratic risk defined as standard deviations of daily residuals is generated by taking the log of variance of the daily residuals: $IdiosyncraticRisk = logvar(\epsilon_{it})$. Next, the daily returns of stocks are computed by taking the difference between log intraday prices: $R_{it} = ln(P_{i,s,t}) − ln(P_{i,0,t})$, in which $P_{i,s,t}$ and $P_{i,0,t}$ are closing and open prices respectively.

3.3. Panel analysis

Panel analysis is used in this research because of its advantages and the disadvantages of using portfolio analysis. First, the traditional analysis of portfolio returns where a portfolio is constructed by sorting stocks according to characteristics, liquidity for example, and then allocated into portfolios, has been more popular recently. This method emphasizes the average portfolio returns to infer the predictive power of the characteristics. If the average portfolio returns are significantly different, the characteristics have the power of prediction. Various stocks, however, have different characteristics which can be used to predict future stock returns. Consequently, the number of portfolios will increase as the number of characteristics grows.

Next, there would be interaction between stocks cross market and over time that could result from some co-movement between corporate stocks, as stated by Mullins (1982). Pooling individual stocks in a portfolio could limit interaction between them, while applying panel data analysis allows individual stocks to interact with each other cross market over time.[3] Finally, the panel data analysis controls for the individual heterogeneity which we would expect to find in this type of data. However, the cross-sectional portfolio and time series analysis do not solve the individual heterogeneity problem.

3.4. Panel data

Daily variables for 150 stocks generated from the intraday sample are set in panel data which comprise 116,583 observations. Selecting new sample data from the original set reduces the size of the panel data to 18,287. The reason for filtering the original sample is that there is considerable dispersion in the number of observations and trades per day among the companies. The new sample is generated by screening out stocks that have less liquidity and avoids the noise caused by substantially short and low liquid stocks. The number of observations for companies varies from 1074 to 1791 instead of from 8 to 1791 after removing the noise stocks. The panel data in the paper is unbalanced due to missing observations for some units during the period 2003–2010. These can be considered as "missing at random", so the fixed effect model suggested by the results of the tests in this paper may not be problematic.[4] An attrition problem for individual stock could be expected to arise in large panel data; for instance, bankruptcy would cause a company to leave the sample. Selecting samples would eliminate this problem. Using panel analysis, attrition causes no problem for the fixed effect model because an attrition problem can be correlated with α_i - the unobserved fixed effect model.[5]

The panel data is summarized in Table 1. In the table, there are three liquidity variables: turnover, nTrades, and BindFreq. nTrades liquidity is defined as the number of trades per day. Turnover liquidity is the number of trades per day in the Norwegian currency (NOK). BindFreq liquidity is the frequency within a trading day at which spreads equal tick size. Idio is

---

[3] In this paper all stocks from the stock market are employed in the panel data.

[4] Suggested by Wooldridge (2013).

[5] This discussion can be found in Advanced Panel Data Methods chapter 14 at http://fmwww.bc.edu/ec-c/s2013/327/EC327.S2013.nn2.pdf.

Table 2

Heteroskedaticity and Autocorrelation Consistent Covariance Matrix – Arellano Method

The consistent covariance matrix generated due to the heteroskedaticity and autocorrelation problem of the fixed effect model examines the relationship between the today returns, and the aggregate risk and liquidity of the previous day. The sample is approximately 18,000 observations over 7 years, from 2003 to 2010. Returns, lag idiosyncratic risk, lag beta, and lag BindFreq are the today returns, previous idiosyncratic risk, beta, and liquidity respectively. The *P* values for significance levels are shown in the note.

	Dependent variable Returns
Lag idiosyncratic risk	$-7.2767e-06^{***}$
	$(1.0502e-06)$
Lag beta	$-1.9530e-06$
	$(3.3736e-06)$
Lag BindFreq	$1.1761e-05^{***}$
	$(2.6275e-06)$

Note:
    * $p < 0.1$.
    ** $p < 0.05$.
    *** $p < 0.01$.

the idiosyncratic risk, lidio is one lag idiosyncratic risk, lbeta is one lag beta, lagBindFreq is one lag BindFreq, and length is the number of observations.

4. Models and empirical results

First, the relationship between returns, risk, and liquidity will be examined. The paper also investigates the correlation between risk and liquidity variables by using single stock panel data.[6]

4.1. Returns, risk, and liquidity relationship

Modelling current returns as a function of one lag idiosyncratic risk, one lag systematic risk, and one lag liquidity is shown in fixed effect model (2)

$$R_{i,t} = b_{i,\beta}\beta_{i,t-1} + b_{i,\epsilon}\sigma_{i,t-1} + LIQ_{i,t-1} + \alpha_i + \nu_{it} \tag{2}$$

where $R_{i,t}$ is the returns of stock i day t, $\beta_{i,(t-1)}$ is the systematic risk of stock i on the previous day, $\sigma_{i,(t-1)}$ is the idiosyncratic risk of stock i on the previous day, and $LIQ_{i,t-1}$ is one-period lag liquidity. In the panel sample, there are three liquidity variables: nTrades, lagBindFreq, and Turnover. lagBindFreq liquidity is chosen for this model because there is a high correlation between the other two liquidity variables and the explanatory variables,[7] that would cause a multicollinearity problem if all these liquidity variables are used in the model. In this model different intercepts – α_i for each individual company are intended to capture individual heterogeneity, which is the different behaviour between individual companies. Individual intercepts are considered as a control of the specific and time-invariant characteristics of individual companies (Hill et al., 2012).

All variables are tested unit root using the Augmented Dickey Fuller test. The results of the test suggest that all variables can be used at level. The research follows the panel data analysis where the pooling, the fixed, and the random effect models should be generated. Testing to choose an appropriate model is then conducted by the Hausman test, the Lagrange test, and the *F* test. The fixed effect model is chosen by the results of the tests.[8] Heteroskedaticity and the autocorrelation consistent covariance matrix following the "Arellano" method are computed due to the heteroskedaticity and autocorrelation problems of the panel fixed effect model. The consistent covariance matrix is presented in Table 2, and the results of the fixed effect model are shown in Table 4 in Appendix 1.[9]

In Table 2, the asterisks denote significance levels; numbers in brackets are standard errors. The P values are given in the note. The estimated coefficients of idiosyncratic risk and of BindFreq liquidity are significantly different from zero, suggested by P values at a significance level of 1%. One-period lag idiosyncratic risk and current returns are in an inverse relation, while the liquidity and current returns have a positive relation. A negative relation between idiosyncratic risk and returns in this

---

[6] Application of a panel of individual stocks for predicting stock returns is discussed by Bauer et al. (2004).

[7] The correlation matrix can be found in Table 9 in Appendix 1.

[8] The results of the fixed effect, the pooling, the random effect model, and the tests can be found in Tables 4–8 in Appendix 1.

[9] This relationship is also examined by employing the whole sample of 116,583 observations; the results of the test are similar to this test in terms of significance. However, the R^2 in the fixed effect model in this case is very small (about 0.001). Filtering the sample raises R^2 to 0.01.

Table 3
Consistent Covariance Matrix for the Risk and Liquidity Model
The model here shows risks on day t, including systematic risk (beta) and idiosyncratic risk, as a function of liquidity variables on day $t-1$. The sample is approximately 18,000 observations over 7 years from 2003 to 2010. The consistent covariance matrix is computed because the fixed effect model has a problem with heteroskedaticity and autocorrelation. The idiosyncratic risk and beta on day t are dependent variables. The liquidity variables on day $t-1$, including nTrades and BindFreq, are independent variables. Numbers in brackets are Std. Error. The P values for significance levels are shown in the note.

	Dependent variables	
	Beta	Idiosyncratic risk
Lag nTrades	0.10404[*]	0.30366[***]
	(0.04903)	(0.05883)
Lag BindFreq	−0.02304	−0.34037[*]
	(0.05367)	(0.13623)

Note:
[*] $p < 0.1$.
[**] $p < 0.05$.
[***] $p < 0.01$.

empirical research is consistent with Ang et al. (2006, 2009). However, the results are different from the empirical work by Spiegel and Wang (2005). They prove that stock returns and idiosyncratic risk have a positive relation, and returns decrease when liquidity is high. In addition, this research does not find a significant effect of one-period lag beta on current stock returns. The flat relationship between returns and beta is consistent with some previous studies. Empirical results found by Hawawini et al. (1989), Ostermark (1991), Corhay et al. (1987), Chan and Chui (1996), Wong and Tan (1991), Cheung and Wong (1992), Ho et al. (2000), Cheung et al. (1993), Lakonishok and Shapiro (1986) and Theriou et al. (2010) suggest that there is either no relationship or an inconsistent one between returns and beta.

This research also addresses the relationship by dividing the panel sample into two periods: 2003–2007 (pre-crisis) and 2008–2010 (crisis period) to separate the effects of the crisis. The results of the test using the two separate periods are similar to the above test using the whole sample. Idiosyncratic risk and the liquidity variable of the previous day affect current returns, and systematic risk (beta) has no relation with the today returns.

The empirical findings show that one lag idiosyncratic risk significantly affects current returns rather than one lag systematic risk. Intuitively, the reasons are perhaps first because of the features of high frequency trading, and second because of the use of single stocks instead of portfolios. High frequency trading means a considerably short time for trading and immediate interaction between specific risk from a company and the flow of markets. In addition, finding an insignificant relationship between systematic risk and returns could be consistent in some ways with investors' expectations. Investors expect that market risk would reduce its impact in algorithmic trading when the Trade News interviewed them to find the reasons for using algorithmic trading. There are many reasons for choosing this type of trading, such as improving trader productivity, reducing market impact, cost, speed, price improvement, and ease of use. 12.4% of investors in 2014 and 11.3% in 2015 chose high frequency trading because of the market impact reduction (The Trade News, 2014, 2015). Hence, these findings contribute greatly to the "idiosyncratic risk pricing returns" literature.

4.2. Risk and liquidity relationship

The data used in this regression is the same as in the previous one. The relationship between risk and liquidity is modelled as in fixed effect model (3)

$$Risk_{i,t} = b_{i,t,} LIQ_{i,t-1} + \alpha_i + \gamma_{it} \tag{3}$$

where $Risk_{i,t}$ is the risk of stock i day t, comprising systematic and idiosyncratic risk, and $LIQ_{i,t-1}$ is past liquidity variables, including nTrades, and BindFreq.

There are three liquidity variables in the panel sample: turnover, nTrades, and BindFreq. However, the turnover liquidity has a high correlation with nTrades and BindFreq. Therefore, it should not be in the regression in order to avoid the multicollinearity problem.

The analysis of this model is similar to that of the previous one, which is to follow the method of panel analysis. Three models, the fixed effect, the random effect, and the polling models, are also generated. The fixed effect model is suggested as an appropriate one from the results of the Hausman, Lagrange, and F tests. Due to heteroskedaticity and serial-correlation problems in the fixed effect model, the consistent covariance matrix is also computed. It is presented in Table 3.

The estimated results in Table 3 suggest that there is a significant relationship between both risks and nTrades liquidity, while BindFreq liquidity has a weak effect on idiosyncratic risk and no effect on beta. Contrary to the empirical work by Malkiel and Xu (2002), in which liquidity and idiosyncratic risk have an insignificant relation, this research proves that both types of liquidity affect idiosyncratic risk, even though BindFreq liquidity has a low effect. In this case, the liquidity variables affect idiosyncratic risk more than the systematic.

5. Discussion

The empirical findings suggest that idiosyncratic risk plays a more pronounced role than systematic risk in asset pricing in high frequency trading because it has a significant effect on returns and is highly affected by liquidity variables, while systematic risk has an insignificant effect on returns and a low effect from liquidity variables.

High frequency trading is expected to improve liquidity, which can be clearly seen by the dramatic increase in the number of trades and it is also confirmed that "AT improves liquidity and enhances the informativeness of quotes" (AT proxy for algorithmic trading), as found in studies by Hendershott et al. (2011). The findings in this research suggest that the number of trades positively affects both idiosyncratic risk and systematic risk. However, idiosyncratic risk is impacted more heavily than systematic risk. The reason is that the coefficient of Lag nTrades (or the number of trades) estimated from the beta regression is lower than that estimated from the idiosyncratic risk regression, at 0.10404 and 0.30366 respectively. In addition, the significance level α of the coefficient of Lag nTrades for systematic risk regression is higher than that for idiosyncratic risk regression, at 10% and 1% respectively (see Table 3). Hence, this could suggest that increases in the number of trades in high frequency trading lead to dramatic effects on both risks, in which the idiosyncratic risk effects are perhaps considerably greater than those of systematic risk.

In addition, high frequency trading from an ecological perspective poses a crucial challenge. According to Doyne Farmer and Skouras (2013), the ecological perspective is an appropriate description for this type of trading because the market has become more systematic and measurable than traditional financial markets, where trading is mainly based on human judgements. In the ecological view, in which a food web is a central notion, algorithmic trading is explained in a similar way to this web. For example, "the food web describes who eats whom. Similarly, in markets a financial food web describes who makes profits from whom" and, "in biology food webs depend on basal species that gather energy either from the sun or…Similarly, financial food webs depend on economic activities that create inefficiencies" (Doyne Farmer and Skouras, 2013). Research by Galaz et al. (2015) also shows the connection between financial markets and ecological changes. They discuss two types of connections between financial markets and ecosystems: direct links (relative to corporate behaviour, financial instruments) attractive to policy makers at the national and international levels, and indirect links. They also provide two diagrams presenting these connections.

The move in financial markets to high frequency trading (or the extension of algorithm trading) has now changed the structure and the nature of the traditional trading market, which before was more based on human interaction. This algorithm market is the most de-personalized of current market forms (MacKenzie, 2014), meaning that this automated trade involves minimal human interaction. Most trading processes are conducted by advances in the sophistication of automated trading technology. An example of the transformation of the nature of financial trade can be seen in changes in the nature of market making. Furse et al. (2011a) state that "the nature of market making has changed, shifting from designated providers to opportunistic traders". Replacing human performance by algorithmic trading raises a further serious problem, because humans can deal with unexpected events much better than computers. Over the long term, if humans are replaced by computers, they will not be available to deal with unexpected events which could be caused by computer trading. Therefore, there could be a problem of out of control technology when markets and most participants in the markets are computer algorithms.

Rapid changes in financial markets could lead to changes in regulations, trading strategies, market perspectives and interaction between agents. Trading in a new algorithm market would bring extremely high speed data processing and dramatically quick data communication from one location to another (in milliseconds). However, critics of high frequency trading indicate that the risk of sudden wide price-swings and the risk of increases in market instabilities could be caused by errant algorithms or incorrect input data (Johnson et al., 2013; Furse et al., 2011b). A report by Blas (2011) in the Financial Times in 2011 entitled "High speed trading blamed for sugar rises" infers that investors struggled to understand a price move because the information on supply and demand had not emerged. The market suffered again with similar price movements several months later. In addition, there was a report in May 2010 on a "flash crash" that had attracted the attention of US regulators on algorithmic trading (Blas, 2011). Therefore, there are fears about high frequency trading, such as out of control technology or finance which could be signals for occurrence of global financial crises.

6. Conclusion

The research first examines how aggregate risk and liquidity affect returns in high frequency trading. Idiosyncratic risk is expected to play a more vital role than systematic risk because of inefficient information, many types of costs, and a short trading time. The empirical results identify specific risk and BindFreq liquidity in one period lag that influence current returns. Systematic risk, however, is not found to impact returns. This evidence strongly supports Merton (1973), Boehme et al. (2009), and other previous studies, that idiosyncratic risk should be priced in cross-sectional returns. Finally, the research investigates the relationship between both idiosyncratic and systematic risk and liquidity. The results of this empirical work suggest that liquidity influences idiosyncratic risk. This influence, however, depends on how liquidity is defined. The number of trades per day in one period lag is more pronounced than the BindFreq liquidity in this research. The results prove that there is a very low significant relationship between liquidity and beta. The two empirical results provide concrete evidence that specific risk plays more important roles than systematic risk in very short interval trading. Moreover,

there could be other crucial challenges such as out of control technology and finance in high frequency trading, to which close attention should be paid.

This study could be extended in several ways. Future research may first work on finding another method to generate idiosyncratic risk instead of the standard deviation of errors estimated from the CAPM or from the three factor model. Finally, it would be also interesting to answer the question of how the effects of idiosyncratic risk could be eliminated though holding a diversified portfolio, not a properly diversified portfolio, the holding of which is impossible.

Appendix A. Returns, risk and liquidity models

Tables 4–9
Fig. 1

Table 4
Fixed Effect Model
The fixed effect model tests the relationship between returns on day t, and aggregate risk and liquidity variables on day $t-1$. The sample is approximately 18,000 observations over 7 years, from 2003 to 2010. Returns, lag idiosyncratic risk, lag beta, and lag BindFreq are the today returns, previous day idiosyncratic risk, beta, and liquidity variables respectively. The P values for the significance levels are shown in the note.

	Dependent variable Returns
Lag idiosyncratic risk	$-7.2767\mathrm{e}{-06}^{***}$
	$(6.7918\mathrm{e}{-07})$
Lag beta	$-1.9530\mathrm{e}{-06}$
	$(1.3821\mathrm{e}{-06})$
Lag BindFreq	$1.1761\mathrm{e}{-05}^{***}$
	$(2.5986\mathrm{e}{-06})$
Observations	18,287
R^2	0.010
Adjusted R^2	0.010
F statistic	58.870^{***} (df = 3; 18,243)

Note:
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Table 5
Pooling Model
The pooling model tests the relationship between returns on day t, and the aggregate risk and liquidity variables on day $t-1$. The sample is approximately 18,000 observations over 7 years, from 2003 to 2010. The pooling model is not chosen as an appropriate model based on the tests shown in Table 7. Returns, lag idiosyncratic risk, lag beta, and lag BindFreq are the today returns, previous day idiosyncratic risk, beta, and liquidity respectively. The P values for the significance levels are shown in the note.

	Dependent variable Returns
Lag idiosyncratic risk	$-5.2290\mathrm{e}{-06}^{***}$
	$(5.5766\mathrm{e}{-07})$
Lag beta	$-7.0627\mathrm{e}{-06}^{***}$
	$(1.1527\mathrm{e}{-06})$
Lag BindFreq	$1.0420\mathrm{e}{-05}^{***}$
	$(2.3741\mathrm{e}{-06})$
Constant	$-7.7525\mathrm{e}{-05}^{***}$
	$(7.2641\mathrm{e}{-06})$
Observations	18,287
R^2	0.010
Adjusted R^2	0.010
F statistic	58.576^{***} (df = 3; 18,283)

Note:
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Table 6
Random Effect Model
The random effect model tests the relationship between returns on day t, and the aggregate risk and liquidity variables on day $t − 1$. The sample is approximately 18,000 observations over 7 years, from 2003 to 2010. The random effect model is not chosen as an appropriate model based on the tests shown in Table 7. Returns, lag idiosyncratic risk, lag beta, and lag BindFreq are the today returns, previous day idiosyncratic risk, beta, and liquidity respectively. The P values for the significance levels are shown in the note.

	Dependent variable Returns
Lag idiosyncratic risk	−7.0801e−06[***]
	(6.7077e−07)
Lag beta	−3.0489e−06[**]
	(1.3726e−06)
Lag BindFreq	1.1345e−05[***]
	(2.5925e−06)
Constant	−9.8959e−05[***]
	(9.6551e−06)
Observations	18,287
R^2	0.010
Adjusted R^2	0.010
F statistic	58.908[***] (df = 3; 18,283)

Note:
[*] $p < 0.1$.
[**] $p < 0.05$.
[***] $p < 0.01$.

Table 7
Results of Tests to Choose an Appropriate Model
The table presents the tests to choose an appropriate model from the three possibilities: the pooling, fixed effect, and random effect models. The Hausman test is used to choose the random effect and the fixed effect. The null hypothesis is that the random effect model is appropriate; the alternative hypothesis is that the fixed effect model is appropriate. The Lagrange test is used to select between the pooling and the random models, in which the null hypothesis is that the pooling model is appropriate, while the fixed effect model is appropriate if the null is rejected. The F test is used to choose the pooling and the fixed effect models; if the null is rejected, the fixed effect model is appropriate. The results of the tests shown in the table suggest that the fixed effect model is the appropriate one.

Test	Null hypothesis	Alternative hypothesis	P-values	Chosen model
Hausman test	Random	Fixed	<2.2e−16	Fixed model
Lagrange test (LM)	Pooling	Random	<2.2e−16	Random
F test	Pooling	Fixed	<2.2e−16	Fixed model

Table 8
Heteroskedaticity and Autocorrelation for The Fixed Effect Model
The table presents the Breusch–Pagan test and Breusch–Godfrey/Wooldridge test for heteroskadaticity and autocorrelation for the chosen fixed effect model. The results suggest that the fixed effect model has a problem with these.

Test	Null hypothesis	Alternative hypothesis	P-values	Comments
Breusch–Pagan test	Homoskedasticity	Heteroskedasticity	<2.2e−16	Heteroskedasticity
Breusch–Godfrey/Wooldridge test	No serial correlation	Serial correlation	<2.254e−12	Serial correlation

Table 9
Correlation Matrix
The table presents the correlation between the variables. The turnover liquidity in one lag period (or laglogturnover) has a high correlation with the other variables: idiosyncratic risk today and on the previous day (or idio and lidio), laglogntrades liquidity, and lagBindFreq liquidity. Avoiding the multicollinearity problem, the turnover liquidity is not chosen as an explanatory variable in this paper.

	Idio	Beta	Lidio	lbeta	Laglogntrades	Returns	Lagreturns	LagBindFreq	Laglogturnover
Idio	1	−0.083	0.816	−0.180	−0.127	−0.075	−0.122	−0.322	−0.455
Beta	−0.083	1	−0.135	0.512	0.325	−0.018	0.007	0.175	0.367
Lidio	0.816	−0.135	1	−0.079	−0.133	−0.082	−0.128	−0.350	−0.538
Lbeta	−0.180	0.512	−0.079	1	0.429	−0.035	0.015	0.149	0.439
Laglogntrades	−0.127	0.325	−0.133	0.429	1	−0.015	0.019	0.111	0.697
Returns	−0.075	−0.018	−0.082	−0.035	−0.015	1	−0.024	0.054	0.042
Lagreturns	−0.122	0.007	−0.128	0.015	0.019	−0.024	1	0.044	0.086
LagBindFreq	−0.322	0.175	−0.350	0.149	0.111	0.054	0.044	1	0.534
Laglogturnover	−0.455	0.367	−0.538	0.439	0.697	0.042	0.086	0.534	1

Series fixed$residuals
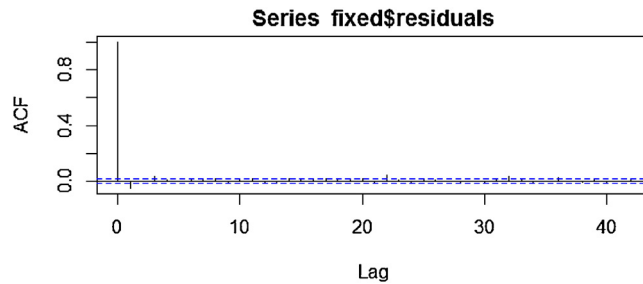

Fig. 1. ACF for the fixed effect model.

Appendix B. Risk and liquidity models

Table 10

Table 10

The Fixed Effect Model for Risk and Liquidity

The table presents the results of the fixed effect model, in which systematic risk (beta), and idiosyncratic risk are modelled as a function of the liquidity variables, Lag number of trades liquidity and BindFreq liquidity. The size of the sample is approximately 18,000 observations over 7 years, from 2003 to 2010. This fixed effect model has a problem with heteroskedaticity and autocorrelation, so a consistent covariance matrix following the "Arellano" method is produced. The results of the consistent covariance matrix and that of the fixed effect model are different. The consistent covariance matrix is chosen to imply the relationship presented in the main part of the paper.

	Dependent variables	
	Beta	Idiosyncratic risk
Lag number of trades	0.10404***	0.30366***
	(0.00600)	(0.01094)
Lag BindFreq	−0.02304*	−0.34037***
	(0.01385)	(0.02526)
Observations	18,287	18,287
R^2	0.017	0.054
Adjusted R^2	0.017	0.054
F statistic	156.998*** (df = 2; 18,244)	525.226*** (df = 2; 18,244)

Note:
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

References

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. J. Finance 61 (1), 259–299.
Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. J. Financ. Econ. 91 (1), 1–23.
Angelidis, T., Tessaromatis, N., 2008. Idiosyncratic volatility and equity returns: UK evidence. Int. Rev. Financ. Anal. 17, 539–556.
Bali, T.G., Cakici, N., 2008. Idiosyncratic volatility and the cross section of expected returns. J. Financ. Quant. Anal. 43 (1), 29–58.
Bauer, R., Pavlov, B., Schotman, P.C., 2004. Panel data models for stock returns: the importance of industries. WP03 8.
Berggrun, L., Edmundo, L., Emilio, C., 2016. Idiosyncratic volatility and stock returns: evidence from the MILA. Res. Int. Bus. Finance 37, 422–434.
Blas, J., 2011, February. High-Speed Trading Blamed for Sugar Rises. Financial Times.
Boehme, R.D., Danielsen, B.R., Kumar, P., Sorescu, S.M., 2009. Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets Miller (1977). J. Financ. Mark. 12 (3), 438–468.
Chan, A., Chui, A.P.L., 1996. An empirical re-examination of the cross-section of expected returns: UK evidence. J. Bus. Finance Acc. 23 (9/10), 1435–1452.
Cheung, Y.-L., Wong, K.-T., 1992. An assessment of risk and return: some empirical findings from the Hong Kong stock exchange. Appl. Financ. Econ. 2, 105–114.
Cheung, Y.-L., Wong, K.-T., Ho, Y.-K., 1993. The pricing of risky asset in two emerging Asian markets – Korean and Taiwan. Appl. Financ. Econ. 3, 315–324.
Corhay, A., Hawawini, G., Michel, P., 1987. The Pricing of Equity on the London Stock Exchange: Seasonality and Size Premium. European Institute for Advanced Studies in Management.
Cremers, M., Halling, M., Weinbaum, D., 2015. Aggregate jump and volatility risk in the cross-section of stock returns. J. Finance LXX, 577–613.
Doyne Farmer, J., Skouras, S., 2013. An ecological perspective on the future of computer trading. Quant. Finance 13 (3), 325–346.
Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. J. Financ. Econ. 91 (1), 24–37.
Furse, C., Haldane, A., Goodhart, C., Cliff, D., Zigrand, J.P., Houstoun, K., Linton, O., Bond, P., 2011a. The future of computer trading in financial markets. Technical report, Final report, Foresight. Government Office for Science, London, UK
http://www.bis.gov.uk/foresight/our-work/projects/current-projects/computer-trading.
Furse, C., Goodhart, C., Cliff, D., Zigrand, J.P., Houstoun, K., Linton, O., Bond, P., 2011b. The future of computer trading in financial markets: An international perspective. Technical report, Final report, Foresight. Government Office for Science, London, UK)
https://www.gov.uk/government/publications/future-of-computer-trading-in-financial-markets-an-international-perspective.
Galaz, V., Gars, J., Moberg, F., Nykvist, B., Repinski, C., 2015. Why ecologists should care about financial markets. Trends Ecol. Evol. 30 (10), 571–580.
Goyal, A., Santa-Clara, P., 2003. Idiosyncratic risk matters! J. Finance 58 (3), 975–1007.

Hatemi-J, A., Irandoust, M., 2011. The dynamic interaction between volatility and returns in the US stock market using leveraged bootstrap simulations. Res. Int. Bus. Finance 25 (3), 329–334.

Hawawini, G., Michel, P., Corhay, A., 1989. A Look at the Validity of the Capital Asset Pricing Model in Light of Equity Market Anomalies: The Case of Belgian Common Stocks. A Reappraisal of the Efficiency of Financial Markets. Springer-Verlag, Berlin, pp. 231–250.

Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? J. Finance 66 (1), 1–33.

Hill, R.C., Griffiths, E.W., Lim, C.G., 2012. Principles of Econometrics, 4th ed. Wiley, Asia.

Ho, Y.W., Strange, R., Piesse, J., 2000. CAPM anomalies and the pricing of equity: evidence from the Hong Kong market. Appl. Econ. 32, 1629–1636.

Huang, W., Liu, Q., Rhee, S.G., Zhang, L., 2009. Return reversals, idiosyncratic risk, and expected returns. Rev. Financ. Stud., http://dx.doi.org/10.1093/rfs/hhp015.

Johnson, N., Zhao, G., Hunsader, E., Qi, H., Johnson, N., Meng, J., Tivnan, B., 2013. Abrupt rise of new machine ecology beyond human response time. Sci. Rep. 3.

Lakonishok, J., Shapiro, A.C., 1986. Systematic risk, total risk and size as determinants of stock market returns. J. Bank. Finance 10 (1), 115–132.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev. Econ. Stat. 47 (1), 13–37.

MacKenzie, D., 2014. A Sociology of Algorithms: High-Frequency Trading and the Shaping of Markets. Edinburgh (Unpublished).

Malkiel, B.G., Xu, Y., 2002. Idiosyncratic risk and security returns. Working paper. University of Texas at Dallas.

Merton, R.C., 1973. An intertempora capital asset pricing model. Econometrica 41 (5), 867–887.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. J. Finance 42 (3), 483–510.

Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 34 (4), 768–783.

Mullins, D.W., 1982. Does the capital asset pricing model work? Harv. Bus. Rev., 105–114.

Ostermark, R., 1991. Empirical evidence on the capital asset pricing model (CAPM) in two Scandinavian stock exchanges. Omega 19, 223–234.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. J. Acc. Econ. 42, 35–52.

Ross, S.A., 1976. The arbitrage theory of capital asset pricing. J. Econ. Theory 13 (3), 341–360.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. J. Finance 19 (3), 425–442.

Spiegel, M.I., Wang, X., 2005. Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk. Yale ICF Working Paper No 05-13; EFA 2005 Moscow Meetings Paper, September 8.

Storesletten, K., Telmer, C.I., Yaron, A., 2007. Asset pricing with idiosyncratic risk and overlapping generations. Rev. Econ. Dyn. 10, 519–548.

The Trade News, 2014. The 2014 Algorithmic Trading Survey – Market Review: Hedge Funds. The Trade News.

The Trade News, 2015. The 2015 Algorithmic Trading Survey-Market Review: Hedge Funds. The Trade News.

Theriou, N.G., Aggelidis, V.P., Maditinos, D.I., Šević, Z., 2010. Testing the relation between beta and returns in the Athens stock exchange. J. Manag. Finance 36 (12), 1043–1056.

Wagner, N., 2004. Time-varying moments, idiosyncratic risk, and an application to hot-issue IPO aftermarket returns. Res. Int. Bus. Finance 18, 59–72.

Wong, K.A., Tan, M.L., 1991. An assessment of risk and return in the Singapore stock market. Appl. Financ. Econ. 1, 11–20.

Wooldridge, M.J., 2013. Introductory Econometrics: A Modern Approach, 5th ed. Cengage Learning, South-Western, USA.