



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Finance

Illiquidity – Measures and Effects

An empirical Analysis of the Nordic Corporate Bond Markets

Christian Engberg and Claes Furthmüller

Supervisor: Jianhua Zhang
Master Degree Project No. 2014:95
Graduate School

Illiquidity – Measures and Effects

An Empirical Analysis of the Nordic Corporate Bond Markets

Christian Engberg & Claes Furthmüller

Supervisor: Jianhua Zhang¹

GM1060 V14 Master Degree Project in Finance



UNIVERSITY OF GOTHENBURG SCHOOL OF BUSINESS, ECONOMICS AND LAW

ABSTRACT

We set out to investigate the microstructure of the Nordic corporate bond markets, especially examining bond illiquidity. The aim was to estimate liquidity premiums on excess yield and determining the most suitable illiquidity measure. This was done through panel data analysis consisting of 1231 bonds within the Nordic markets. We considered two models used in previous literature and concluded that ‘Model I’ had a higher explanatory power. We found evidence of possible liquidity premiums of 58.2 bps and concluded that the negative autocovariance of the relative price changes of a bond outperformed the other illiquidity proxies used.

¹ We would like to thank our supervisor for his appreciated input and comments improving our thesis.

1. INTRODUCTION AND RESEARCH QUESTIONS	3
1.2 HYPOTHESIS	3
1.3 STRUCTURE	4
1.4 LITERATURE REVIEW:	4
2. BACKGROUND AND THEORY	8
2.1 ECONOMETRIC MODELS (PANEL DATA MODELS)	8
2.1.1 FIXED EFFECTS:	9
2.2.2 RANDOM EFFECTS	9
3. METHOD	12
3.1 'MODEL I'	12
3.2 'MODEL II'	13
3.3 MARKET FACTORS	13
3.3.1 MARKET RISK	13
3.3.2 CREDIT RISK	14
3.3.3 TERM FACTOR, DEFAULT FACTOR AND THE FAMA & FRENCH TWO-FACTOR MODEL	14
3.4 AUGMENTATIONS	15
3.4.1 DURATION	15
3.4.2 CREDIT RATING	16
3.5 LIQUIDITY MEASURES	16
3.5.1 Γ	16
3.5.2 REPORTED BID-ASK SPREAD	17
3.5.3 ZEROS/MISSING PRICES	17
3.5.4 AGE	17
3.5.5 AMOUNT ISSUED	18
3.5.6 COUPON	18
3.6 ECONOMETRIC METHODOLOGY	19
4. DATA AND ANALYSIS	21
5. DISCUSSION AND CONCLUSIONS	31
5.1.1 SUGGESTIONS FOR FURTHER RESEARCH	32
WORKS CITED	34
APPENDIX 1	36

1. Introduction and research questions

“Liquidity refers to the ease and rapidity with which assets can be converted into cash”

- (Hillier, Ross, Westerfield, Jaffe, & Bradford, 2010)

We write this thesis as a part of our Masters degree in Finance at the School of business and law at the university of Gothenburg. We investigate the microstructure of the corporate bond market that includes the Nordic company bonds listed on different exchanges. We also examine how the liquidity affects the yield spread over the benchmark with different measures of liquidity.

Our goal is to measure and analyze liquidity effects on the Nordic corporate bond market and to determine what or which measures and model that provide the best fit.

One of the reasons for examining the liquidity in a poorer lit market is due to adverse selection. In a less transparent market, the two exchanging parties are more likely to have larger differences in perception of the fair price. As a consequence the bid-ask could increase together with the yield spread. This friction increases the cost of trading hence decreasing liquidity in the market (Hillier, Ross, Westerfield, Jaffe, & Bradford, 2010).

For our study we use a dataset consisting of panel data with prices, yields and other bond characteristics for some 2805 bonds in the Nordic market in between 2009 and 2014.

1.2 Hypothesis

- Our first hypothesis is that liquidity is priced in the Nordic corporate bond market.

Investors are surely concerned with the ability to liquidate their assets and we believe that the Nordic corporate bond markets are no exceptions and we test this hypothesis by examining whether or not our illiquidity proxies are statistically significantly different from zero.

- The second hypothesis is that our main choice of illiquidity measure, γ , the negative autocovariance in a bonds price is the liquidity measure with the highest explanatory power compared to the other measures included in this paper.

The measure introduced by Bao Pan and Wang (2011) should capture the transitory price effect caused by illiquidity and thereby some of the properties of illiquidity. It is a direct

measure that has been shown to have large explanatory power in earlier research. We will evaluate this hypothesis both economically and statistically.

1.3 Structure

In section 1.4 we will go through some of the existing literature on the subject, in the following section we will introduce the theories and models linked to our subject. Then we will introduce our data set more thoroughly and discuss our choice of method and models in section 3. Thereafter we will present our results and analyze them in section 4 after which we will draw our conclusion and summarize with a discussion of our findings in section 5.

1.4 Literature Review:

In this section we will present the reader to some of the most influential articles and papers concerning the subject of liquidity in bond markets. We will briefly discuss the articles that are most related to our research but we sincerely recommend every genuinely interested reader to study any or all of the cited work for further understanding of the subject.

In their 2011 journal article Bao, Pan & Wang aim to measure the level and impact of the illiquidity in the US corporate bond market as well as to study the behavior of illiquidity. They establish a strong link between illiquidity in the US bond market and bond prices in their data set of transaction data between 2003 and 2009. They find that illiquidity is the most important factor for high rated bond yield spreads. Their liquidity measure, γ , dominates their proxy for credit risk, a credit default swap (CDS) Index and they find no additional explanatory power in their market risk proxy, the Chicago board Options Exchange Volatility Index (CBOE VIX).

When constructing their liquidity measure they are making use of the illiquidity feature of transitory price movements. Earlier research such that by Niederhoffer and Osbourne (1966), Grossman and Miller (1988), Huang and Wang (2009) amongst others have established a relationship between negative serial covariation and illiquidity.

They begin their reasoning by defining the clean price of a bond as P_t and $p_t = \ln P_t$ and then they assume that p_t consists of f_t , its fundamental value that follow a random walk and u_t as the impact of illiquidity. This u_t is believed to be temporary and using that they define γ as the negative of the autocovariance in relative price changes:

$$\gamma = -Cov(\Delta p_t, \Delta p_{t+1}). \quad (1)$$

$$\Delta p_t = \ln\left(\frac{p_t}{p_{t-1}}\right). \quad (2)$$

By doing so they claim to get a direct measurement of liquidity irrespective of any asset pricing models. When they perform monthly cross-sectional regressions they find that γ outperforms all of the other suggested proxies such as effective bid-ask spread and the percent of zero returns.

In the Fama & French (1993) article they investigate the risk factors determining the variation in returns. The authors use their previous three-factor model as a starting point for their analysis, which was published in the Journal of Finance 1992. Their three factors constituted of one factor that they refer to as the “overall market factor”, one factor for “firm size” and last one factor for “book-to-market equity”.

Eugene Fama and Kenneth French then proceed to develop a two-factor bond model in which they incorporate two “bond-market factors”. These two factors are to explain the risks associated interest rate risk, *TERM*, and default risk, *DEF*.

Throughout their study the authors construct these factors based on the previous work by Chen, Ross and Roll (1986). They chose to define *TERM* as “the difference between the monthly long-term government bond return and the one month Treasury bill rate measured at the end of the previous month”. The factor *DEF* is defined in this way “the difference between the market return on a market portfolio of long-term corporate bonds and the long-term government bond return”. One thing that differs Fama & French and Chen, Ross and Roll is that Fama & French do not limit themselves to cross-sectional analysis.

In a paper by Howeling, Mentink and Vorst (2003) they aim to both measure liquidity in European corporate bonds as well as examine which liquidity measure has the greatest explanatory power. The authors apply a modified version of Brennan and Subrahmanyam’s (1996) approach that in turn took use of Fama and French (1993) three-factor model. The authors modify this model by using yield-to-maturity instead of realized returns and switching to Fama and French (1993) two-factor bond market model controlling for the term structure and the credit risk of the bond market. They also add 3 bond characteristics; denomination currency, duration and the bond rating to fully explain the risk premiums connected to term and default. The third modification to the Brennan and Subrahmanyam’s method is that

instead of using Kyles' (1985) direct measure of liquidity they test eight indirect measures; Issued Amount, Coupon, Listed, Age, Missing Prices, Price volatility, Number of Contributors and Yield Dispersion.

They find that 7 out of the 8 suggested measures of liquidity are statistical significant where age and yield dispersion are linked to the highest premiums found but price volatility and number of contributors seems to have the greatest explanatory power.

McCulloch (1975) examines the liquidity premium in the US market in his article. He chooses to define the liquidity premium in government bonds as "the difference between a forward interest rate and the market's expectation of the corresponding future spot rate." For his paper, he chose the Cagan-Roll method to estimate the 'post-Accord' liquidity.

His methodology is to divide their sample into four different time periods based on different American administrations. The author observes that there seems to be a different mean and/or variance between the sample periods and therefore decides to test five different hypothesis against a likelihood ratio test. Through these tests of the hypothesis conducted in the study, he conclude that he cannot reject the fourth hypothesis, meaning that there has been a quite stable post-Accord premium although the variance of the forecasting errors has not been constant.

The author concludes that there is a liquidity premium that is significantly greater than zero. McCulloch also says that he cannot conclude that there has been any variance in the post accord mean of the liquidity premium between the time periods.

Gopalan, Kadan and Pevzner (2012) investigate if there is a correlation between the liquidity of the assets in a firm's balance sheet and the liquidity of the firms stock.

Their definition of a liquid asset is "an asset is liquid if it can be converted into cash quickly and at a low cost" and in the article the authors uses four different measures for liquidity:

1. Amihud's Liquidity Measure:

$$ILLIQ_{i,t} = \frac{1}{N_{i,t}} \sum \sqrt{\frac{|R_{i,j}|}{VOL_{i,j}P_{i,j-1}}}, \quad (3)$$

where $R_{i,j}$ is the return of stock i on day j, $VOL_{i,j}$ is the volume sold of stock i on day j measured as millions of shares and $P_{i,j-1}$ is the closing price of stock i on the previous day.

The measure should then capture price impact of sold quantities.

2. Implicit bid-ask spread, s , by Roll:

$$s_{i,t} = \sqrt{-cov(R_{i,j}, R_{i,j-1})}. \quad (4)$$

Another measure based on the negative autocovariance of the prices of an asset. This proxy for bid-ask spreads has been proven to explain ½ of the reported spread.

3. Annual average effective bid-ask spread:

The authors calculate this measure from intraday transaction data and define it as the “ratio of the absolute difference between the trade price and the midpoint of the associated quote and trade price”.

4. Pastor – Stambaugh Gamma:

The idea behind the measure is that the prices of highly illiquid assets are probably sensitive to large trades. Pastor and Stambaugh (2003) proposed regressing stock returns on lagged returns and lagged volumes traded. The coefficient of the trade volumes is used as a measure of the illiquidity. They conclude that high liquidity in the balance sheet of firms has a positive relationship to liquidity for that firms stock.

2. Background and Theory

In this section we will present the background to our subject, its subtopics as well as the financial and econometrical theories that lay the foundation for our study.

Liquidity, or illiquidity, as a phenomenon has been known since the days of Adam Smith (Palyi, 1936). Since then, much research has gauged its effects, determinants and implication for asset pricing. The definitions may differ slightly but we use the opening quote of this paper as our definition. The measurement of liquidity however is more problematic since it cannot be directly observed without in fact selling the asset for which you want to know the liquidity. We will consider several direct and indirect measures of illiquidity and try to establish some economical reasoning to them and then evaluate them econometrically.

2.1 Econometric Models (*Panel data Models*)

Baltagi (2008) describes the difference between Panel data regression, time-series regression and cross-sectional regression as that in panel data, the variables has a double subscript. In time-series, we have a subscript for time, denoting the different time periods. In cross-sectional regressions we have a subscript describing the different entities and in panel data we have both a subscript for time as well as a subscript for the different entities.

A stylized example of panel data can be seen below in table 1:

i	t	x_1	x_2
1	2007	2	2.5
1	2008	4	2.6
2	2007	3	3
2	2008	6	4.5
3	2007	4	2
3	2008	8	2.5

Table 1: Table representing a stylized example of Panel Data

As can be seen in the table above, we have observations for the different entities for different time periods.

Table (1) is also what is called balanced panel data, this is since the observations for each entity is equal, that is that for every entity in table (1), we have two observations. It is called

un-balanced panel data if the number of observations for each entity is different, for example if we would have two observations through time for some entities and only one for others.

In the following sections for the fixed- and random effects we will show how these two models as described by Wooldridge (2009) to make it clearer for the reader what the differences are between the two models.

2.1.1 Fixed Effects:

Wooldridge (2009) explains that the fixed effects estimation aims to exclude the unobserved effect (α_i) by time-demeaning the data. Below you can see our starting model before time-demeaning our data in the fixed effects transformation for one variable:

$$y_{i,t} = \beta_1 x_{i,t} + \alpha_i + u_{i,t}. \quad (5)$$

If we consider just the average time effects of (1), we arrive at the following equation:

$$\bar{y}_i = \beta_1 \bar{x}_i + \alpha_i + u_i. \quad (6)$$

The last step in the fixed effects transformation is to subtract (5) from (6) and by doing that, we exclude the unobserved effect α_i since it is assumed to be fixed across time.

$$y_{i,t} - \bar{y}_i = \dot{y}_{i,t} = \beta_1 \ddot{x}_{i,t} + \ddot{u}_{i,t}. \quad (7)$$

If we extend this model by adding k-different independent variables, our model then becomes:

$$\dot{y}_{i,t} = \beta_1 \ddot{x}_{i,t1} + \beta_2 \ddot{x}_{i,t2} + \dots + \beta_k \ddot{x}_{i,tk} + \ddot{u}_{i,t}, \quad t = 1, 2 \dots T. \quad (8)$$

The time-demeaned model is then estimated through a pooled OLS procedure.

The fixed effects models, since they are time-demeaning the equation with the unobserved effects (5), it excludes all of the time-invariant variables that might be in the sample.

2.2.2 Random Effects

With the random effects model, we start with the same “unobserved effects model” as in the case with the fixed effects, i.e:

$$y_{i,t} = \beta_0 + \beta_1 x_{i,t1} + \dots + \beta_k x_{i,tk} + \alpha_i + u_{i,t}. \quad (9)$$

One of the main differences between the random effects model and the fixed effects model is the view of α_i . In the fixed effects model, it is believed to be correlated with the with at least

one of the independent variables while the random effects model imposes the following condition for α_i :

$$Cov(\alpha_i, x_{itj}) = 0 \forall t, j. \quad (10)$$

To quote Wooldridge (2009) “the ideal random effects assumptions include all of the fixed effects assumptions plus the additional requirement that α_i is independent of all explanatory variables in all time periods”.

As can be understood from the quote above, the random effects model should only be used when the covariance between the explanatory variable and α_i is zero.

The aim of this model is the same as in the fixed effects model, i.e. that we want to get rid of α_i . Since we now believe that α_i is not fixed over entities, we cannot simple time-demean the unobserved effects model as with the fixed effects model but instead we have to make some transformations.

The first step is to construct an error term, consisting of both α_i and u_{it} , which will give us a “composite error term” which would equal to:

$$v_{i,t} = \alpha_i + u_{i,t}. \quad (11)$$

Since the creation of v_{it} , we now know that α_i is present in every point in time and v_{it} is correlated in time, we can make the following statement:

$$Corr(v_{it}, v_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} \forall t \neq s. \quad (12)$$

Wooldridge then further proceeds by describing a new variable, λ :

$$\lambda = 1 - \sqrt{\left[\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2} \right]}. \quad (13)$$

Using the λ defined in (9) and time average of y . We have now arrived at the random effects model with “quasi-demeaned data”:

$$\begin{aligned} y_{it} - \lambda \bar{y}_i &= y^{re} = \\ &= \beta_0(1 - \lambda) + \beta_1(x_{i,t1} - \lambda \bar{x}_{i,1}) + \dots + \beta_k(x_{i,tk} - \lambda \bar{x}_{i,k}) + (v_{i,t} - \lambda \bar{v}_i). \end{aligned} \quad (14)$$

One major benefit that stems from using random effects is that it allows for time invariant variables. (ibid)

To decide whether to use the fixed effects routine or the random effects one, researchers are known to apply the work by Hausman (1978) in order to clarify what model to use.

When applying the Hausman test to decide what model to use, the reasoning is that one should use the random effects approach when we cannot reject H_0 . The null hypothesis of the test is that the Random Effects is more efficient than Fixed Effects. Rejecting the null hypothesis would then imply that one should use the fixed effects model since it needs fewer assumptions.

Our next test was designed by Breusch and Pagan (1980) in order to determine if the more suitable model to use is the Random Effects model or if one should use pooled OLS. The LM-test for Random Effects evaluates whether or not the variance of u is zero. If the variance of u is zero, then we do not have any random effects and therefore the model to be preferred is the pooled OLS.

3. Method

In this section we will develop the methodology followed in this study.

To collect and organize our data into an unbalanced panel we used Microsoft Excel and the Bloomberg add-in. All econometric tests and analysis was conducted in Stata 12.

Our choice of method is closer to that of Bao, Pan and Wang, however we incorporate some of the other suggested liquidity measurements in the literature. Other than γ we also regress the excess yield on measures such as the reported bid-ask spread, the number of non-trading days for each bond, both suggested by Goplan et al, among with some other proxies further explained below. We also chose to incorporate some of the market factors such as the modified Fama and French two-factor bond model and the proxies suggested by Bao, Pan & Wang to properly examine the interest rate and credit risks premiums.

From our original data set we needed to omit bonds that were convertible or putable and bonds that did not trade at least 25% of the days in our sample. This is done in accordance with Bao, Pan & Wangs' study.

We begun by evaluating two rather similar models, the one suggested by Bao et al and tested it against the augmented Fama and French two-factor model as used by Howeling et al. Both models aim to describe the yield of a bond based on some market factors, some bond specific characteristics and one ore more liquidity measurements. For the dominant model we will then measure the fit without any liquidity measure and then add our liquidity measurements for comparison. We tested the different measures against each other by following the methodology of Goldreich, Hanke and Nathy (2005) where they rank their different proxies by pairwise regressions with all possible combinations. They compare how often a proxy adds explanatory power to the regression and the number of times a measure dominates another.

We chose to fit our data using the random effects model; this conclusion is based on the Hausman-Test as well as the Breusch and Pagan Lagrange Multiplier Test for Random Effects. The specifications were also tested for heteroskedasticity and autocorrelation through the methodology described in the theory section.

3.1 *'Model I'*

The Bao, Pan & Wang basic model is defined as below and will be referred to as 'Model I'.

$$Y_{i,t} = \beta_0 + \beta_1 cdsindex_{i,t} + \beta_2 volindex_{i,t} + \beta_3 duration_{i,t} + \beta_4 credit\ rating_i + \alpha_i + u_{i,t}. \quad (15)$$

We extended this basic model by adding our different measures for illiquidity

$$Y_{i,t} = \beta_0 + \beta_1 cdsindex_{i,t} + \beta_2 volindex_{i,t} + \beta_3 duration_{i,t} + \beta_4 credit\ rating_i + \beta_5 L_{i,tk} + \alpha_i + u_{i,t}. \quad (16)$$

Where $L_{i,tk}$ is the k:th measurement of illiquidity at time t . As mentioned above we performed pairwise regressions including two measures at each time:

$$Y_{i,t} = \beta_0 + \beta_1 cdsindex_{i,t} + \beta_2 volindex_{i,t} + \beta_3 duration_{i,t} + \beta_4 credit\ rating_i + \beta_5 L_{i,tk} + \beta_6 L_{i,tj} + \alpha_i + u_{i,t}. \quad (17)$$

Where $L_{i,tj}$ is the j:th illiquidity measure at time t for bond $i \forall j \neq k$.

The subsequent ranking was performed by assigning one point to the illiquidity measure in question if that measure dominated the other at a 5% significance level. We then summed the points for each proxy in order to evaluate the econometrical explanatory strength.

3.2 'Model II'

The second model evaluated is based on the model used by Houweling, Mentink and Vorst (2003), which is an extended version Fama-French two-factor bond model. In our thesis this will be referred to as 'Model II' and is defined as:

$$Y_{i,t} = \beta_1 termfactor_{i,t} + \beta_2 deffactor_{i,t} + \beta_3 duration_{i,t} + \beta_4 credit\ rating_i + \beta_5 L_{i,tk} + \alpha_i + u_{i,t}. \quad (18)$$

This model was estimated through pooled OLS in the paper by Houweling, Mentink and Vorst but the F-test for the regression suggested that we should be using fixed effects instead. We also performed the Hausman test that also favored the fixed effects model compared to the random effects. Otherwise the basic approach is the same as in 'Model I', i.e. that we extend the model by including different illiquidity measures.

Below follows introduction and explanation for the four market factors, the bond characteristics as well as all liquidity measures.

3.3 Market factors

3.3.1 Market risk

In the now classic Capital Asset Pricing Model (CAPM) an assets return is dependent on its co-movement with the market. Fama & French extended the model to include some firm

specific characteristics and also made corresponding model for bonds. In that model they included a market risk factor and an interest rate factor described further below.

In the paper by Bao, Pan & Wang they use the CBOE VIX to proxy for market risk. Similarly we use VIX to get the inherent risk in the markets we explore. The variable is quoted as integers i.e. if at a specific day the market volatility was 20% it will be quoted as 20. This is important to remember when analyzing our results later.

3.3.2 Credit risk

A bondholder must be concerned with the issuers' ability to pay coupons as well as the face value. Hence the inherent credit risk in the market is of importance when pricing bonds. To capture the market credit risk Bao, Pan & Wang suggest using a credit default swap (CDS) Index as a proxy. We adapted this methodology and used iTraxx SOVX West Europe as proxy for the market risk in the Nordic bond market. To our knowledge there exist no CDS indices for each Nordic corporate bond market hence our choice of proxy.

3.3.3 Term factor, Default factor and the Fama & French two-factor model

The two famous Economists Eugene Fama and Kenneth French (1993) suggested in their article that two factors could be used to explain the differences in bond returns.

The two factors chosen by Fama and French were *TERM* and *DEF*. The variable *TERM* is meant to be a proxy for the interest rate risk. This was in their article defined as the difference between the long-term government bond monthly return and the short-term government Treasury bill rate at the end of the previous month.

The other factor described in the article, *DEF*, is defined as the difference between the return of a portfolio of long-term corporate bonds and the long-term government bond. Both of these factors are inspired by the article written by Chen, Roll and Ross (1986).

We have made some modifications to these factors. First of all, we are regressing against yields instead of returns as suggested by Amihud and Mendelson (1991). Likewise we use the yields in both the two Fama and French factors to better proxy for expected bond returns.

The two factors are defined as:

$$TERM_{i,t} = Y_{c,t}^{LT} - Y_{c,t}^{ST}. \quad (19)$$

$Y_{c,t}^{LT}$ = Yield for a country c 's long – term government bond at time t .

$Y_{c,t}^{ST}$ = Yield for a country c 's short – term government bond at time t .

$$DEF_{i,t} = Y_{c,t}^{LT Corp} - Y_{c,t}^{LT Gov}. \quad (20)$$

$Y_{c,t}^{LT Corp}$ = Return on a market portfolio of corporate bonds in country c at time t .

$Y_{c,t}^{LT Gov}$ = Return on a long term government bond in country c at time t .

As a proxy for $Y_{c,t}^{LT Corp}$ we used different indices depending on the origin of the company. For the Swedish and the Norwegian market we found specific indices namely, BSEK (Bloomberg SEK Investment Grade Scandinavian Corporate Bond Index) and BNOK (Bloomberg NOK Investment Grade Scandinavian Corporate Bond Index). For Finland and Denmark, we could not find country specific indices so we used BSCA (Bloomberg Investment Grade Scandinavian Corporate Bond Index).

3.4 Augmentations

In addition to the Fama and French-factors several other authors have tried to further examine the microstructure of the corporate bond market. Gebhardt, Hvidkjaer & Swaminathan (2005) and Houweling, Mentink & Vorst (2003) both included several bond specific characteristics to further add explanatory power and widen the understanding of the bond market microstructure. We chose to include duration and credit rating.

3.4.1 Duration

“The duration of a bond...is a measure of how long on average the holder of the bond has to wait before receiving cash payments” (Hull, 2012)

Hull defines duration as:

$$D = \frac{\sum_{i=1}^n t_i c_i e^{-y t_i}}{B}, \quad (21)$$

where:

t_i = time i

c_i = cashflow i

y = bond yield

B = bond price

(ibid.).

We added the duration for all bonds since it will further contribute to explain the interest rate risk connected to that bond.

3.4.2 Credit Rating

To further assess some default premium to the individual bonds we chose to include the credit rating as a proxy for their credit worthiness. Since some of the bonds had credit ratings from different institutes, and since we wanted the credit ratings in numerical form rather than in letters, we chose to create our own composite credit rating. This was done through the help of the conversion table done by International Bank of Settlements (2014). We started by constructing a conversion table as can be seen in table A1 in appendix 1. Since Moody's have a different system than S&P and Fitch, we converted the credit ratings that were performed by Moody's into the same scale as the other two and then assigned them numerical values.

After that we created a binary variable for investment-grade bonds, which in this case would be bonds that has a rating above 12. We also create a second binary variable for this purpose, except that it now only includes BBB- to AAA rated bonds.

Lastly we augment the model with our different liquidity measurements to examine if any add explanatory power to the model and if we could find one superior proxy.

3.5 Liquidity Measures

Below we will briefly explain and discuss the different measures liquidity.

3.5.1 γ

As described in the literature review section Bao, Pan and Wang develop a direct measure of illiquidity, γ , that is defined as the negative of the autocovariance of the returns or more formally:

$$\gamma = -Cov(\Delta p_t, \Delta p_{t+1}). \quad (22)$$

It is based on some of the properties of illiquidity; that it is transitory and that it arises from market friction. The measure should be handled and interpreted with some caution, the underlying assumption that all deviations from the bonds fundamental value stem from illiquidity might be heroic. It does however capture some properties of the price movements and have been proven significant in earlier research and as such we will at least examine and evaluate it as an illiquidity measure. We thus expect it to have a positive effect on the yield spread. The difference between this measure of liquidity and the one that Roll (1984)

developed is that this measure allows for positive autocovariance. The implicit bid-ask spread takes the square root of the γ described above and therefore omits any positive autocovariance in returns.

3.5.2 Reported bid-ask spread

Another way to measure liquidity is to make use of market friction e.g. the bid-ask spread as reported by Bloomberg. The bid-ask spread is simply the difference from the last quoted ask price and the last quoted bid price:

$$s_{i,t} = P_{i,t}^{ask} - P_{i,t}^{bid}. \quad (23)$$

If this measure could be a proxy for illiquidity we should find a positive relationship with the yield spread.

3.5.3 Zeros/Missing prices

As can be understood from the definition of liquidity quoted in the first section of this thesis a liquid asset is one that can be sold easily. One way of measuring the easy of selling an asset could be to examine the frequency of trades and trading volumes for that asset. Since we have no intra daily data or any data on the traded volume we proxy this by the lack of price movements or the lack of price data. Different types of this measure is widely used in the field but we chose to define it as the percentage number of days where the price has not changed for each individual bond:

$$prc\ zero_i = \frac{\text{Number of non trading days for bond } i}{\text{Total number of days for bond } i}. \quad (24)$$

We suspect that this measure should have a positive relationship with the yield spread. As suggested by Bao, Pan & Wang (2011) we removed any bond that had a value exceeding 75%.

3.5.4 Age

In the light of the findings of Sarig & Warga (1989) where they discovered that as a bond gets older a larger portion of the total amount issued end up in low frequency trading portfolios, a bonds' age could be linked to liquidity. We would then expect the binary variable to have a positive effect on the yield spread. We specified an "old" bond as a bond that was more than two years old.

3.5.5 Amount issued

As we lack volumes data we also incorporate an indirect measure of volume and hence illiquidity. First proposed by Fisher (1959) the amount issued is strongly and positively correlated with traded volumes. The reasoning goes that information cost of a larger issue is less than that of a smaller issue and that the transaction cost then decreases.

This measure should then have a negative relationship with the yield spread since larger issuances would mean lower information cost and lower spread. To scale the measure we took the natural logarithm of the issued amount.

3.5.6 Coupon

The level of coupon of a bond can also be seen as a measure of liquidity. Consider two equal bonds that only differ in liquidity. For investors to choose the less liquid bond the issuer would need to compensate by providing larger cash flows, however it might be a noisy measurement since bonds with higher probability of default probably also would be required to give larger coupons. Another aspect is that a higher coupon will ensure the bearer of the bond to larger cash flows thus increasing the speed at which she will recoup her investment. This might make the bond more attractive and thus lowering the spread from the benchmark. As stated, it is surely a noisy measure but nonetheless we will regard it as an illiquidity measure. Our hypothesis is that higher coupons should have a positive sign in our regressions but close to zero. Below in table (2) we have summarized our ex-ante expectations of our different illiquidity proxies:

Ex-ante expectations for our proxies	
Variables	Expected sign
Y	+
Bid-Ask spread	+
% zero/Missing prices	+
Age	+
Amount issued	-
coupon	+

Table 2: Table depicting a summary of our expectations for the proxies introduced above

3.6 Econometric methodology

As for our econometrical models we refer to the theory described in section 2. In this section we focus on specific tests and correction for several issues in time series analysis and in statistics in general.

Heteroskedasticity is defined as inconsistency in the variance of the error-terms. This will render in the estimates not being efficient (Newbold, Carlson, & Thorne, 2010).

To test for heteroskedasticity within the panel data setting we follow the procedure proposed by Wiggins and Poi (2013)².

We begin by forcing the model to correct for heteroskedasticity using a model that produces ML-estimators and then store the likelihood. The second step is to fit the model without forcing for heteroskedasticity and storing the likelihood estimates.

When this is done, we perform a likelihood test between the two fitted models where the null hypothesis is that the error terms are distributed homoskedastically. We can reject this hypothesis for all of our models.

We also wanted to test for the presence of autocorrelation, this we did through the built-in function *xtserial* in STATA, which is based on Wooldrige (2002) and Drukker (2003).

This method makes use of the first difference method, which yields the following equation:

$$y_{i,t} - y_{i,t-1} = (X_{i,t} - X_{i,t-1})\beta_1 + \epsilon_{i,t} - \epsilon_{i,t-1}. \quad (25)$$

Wooldridge's procedure and conclusion is that if:

$$\text{Corr}(\Delta\epsilon_{i,t}, \Delta\epsilon_{i,t-1}) = -0.5, \quad (26)$$

then we do not exhibit serial correlation, and this test is incorporated in the statistical software STATA under the command *xtserial*. The null hypothesis in this test is that we do not exhibit any first-order correlation.

After proceeding with these tests, we could conclude that we had presence of both heteroskedasticity as well as autocorrelation. The next step was then to correct for both of these issues since we wanted to have efficient and consistent estimators of our variables.

² The results for our tests are visible in Appendix 2 for 'Model I' and Appendix 3 for 'Model II'

The approach that we chose to use to correct for these issues was to use clustered standard errors as suggested by Hoechle (2007).

To further examine if we had autocorrelation after conducting our corrections, we performed a Durbin-Watson test our regressions (Durbin & Watson, 1950).

After performing the corrections for both autocorrelation as well as for heteroskedasticity we performed an extension of the Hausman Test that allows us to compare the random effects model with our corrected standard errors with the corresponding fixed effects model. This test has the same interpretation as the basic Hausman test, i.e. the null hypothesis is that the random effects model is the model to prefer.

A further step to assure that the random effects model is the that we should use was to perform another post estimation test, the Breusch and Pagan Lagrange Multiplier test for Random Effects.

Through the usage of the Hausman-test, The F-test of the coefficients and the Breusch Pagan LM-test we could conclude that we should use the random effects model for 'Model I' and the fixed effects model for our 'Model II'. Since we believe that our time-invariant variable *creditrating* has a large impact on the yield spread we chose to create interaction terms for that variable when using the fixed effects model. We then also applied the same methodology for the time-invariant illiquidity measures.

We were also interested in the difference between investment-grade bonds and speculative-grade bonds. We created a binary variable that separated the investment-grade bonds and the speculative-grade bonds and then used 'Model I' to examine the differences.

4.Data and analysis

In this section we will present the results from our various regressions, some statistics for our variables of interest and interpret the above stated results.

Our dataset consists of daily bond prices, yield spreads, and a multitude of relevant variables explained in the previous section, for all currently traded corporate bonds for companies based in Sweden, Denmark, Finland and Norway. The set span between 2006-01-01 until 2014-02-06 and consists of 2805 bonds and a total of 908694 observations. All bonds have a fixed coupon and any convertible bonds were omitted. The benchmark risk free proxies were chosen based on the currency of the bond and we used the related government bond with similar maturity and coupon when possible.

After some modifications and after clearing the data according to what is mentioned in the method section the data set span between 2009-01-01 to 2014-02-06 and consist of 1231 bonds and a total of 597673 observations.

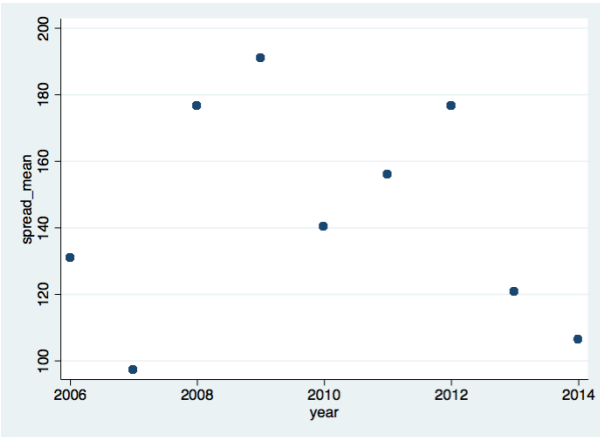


Figure 1: Mean Spread from Benchmark by Year in Scandinavia

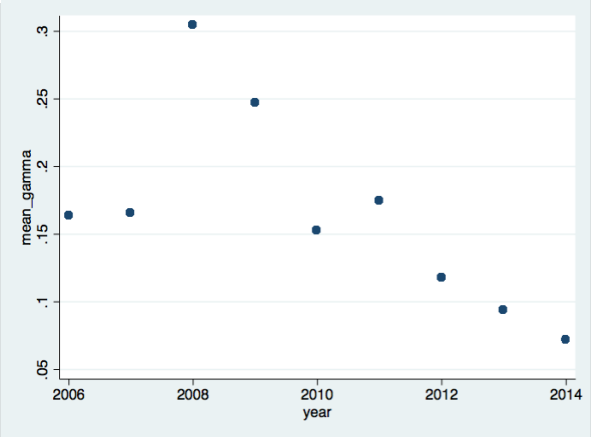


Figure 2: Mean γ per Year in Scandinavia

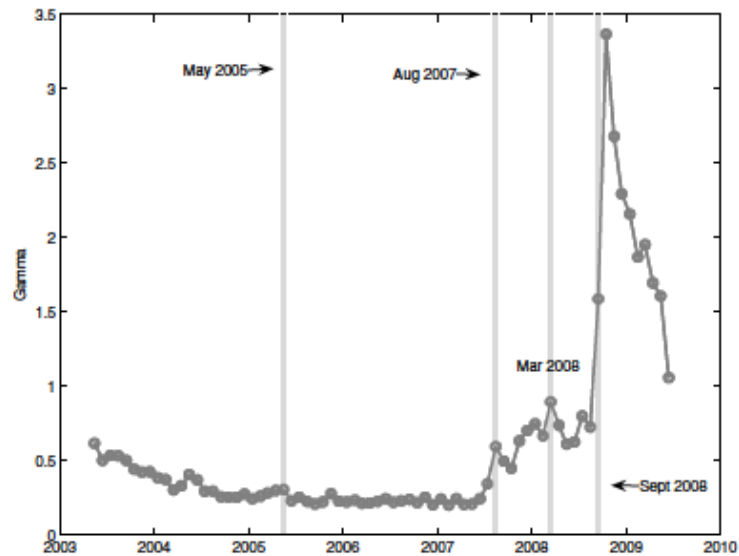


Figure 3: portraying the central tendency of γ in the US market over time. Taken from “The Illiquidity of Corporate Bonds” (Bao, Pan, & Wang, 2011)

As can be seen when comparing figure (2) to figure (3) there is a spike in illiquidity during the financial crisis. The magnitude of γ is clearly larger in the US market for the entire comparable period. In figure (1) the aggregate mean yield spread from benchmark is plotted over time with clear spikes during the two financial turmoils included in the period.

	Variable characteristics			
	Mean	Std dev	Min	Max
<i>cds index</i>	155,8037	94,3146	46,875	386,667
<i>vol index</i>	22,8929	6,8346	13,8195	56,3255
<i>duration</i>	4,8938	3,0022	0,01	32,03
<i>credit rating</i>	18,2566	3,6281	3	21
<i>% zero</i>	0,1085	0,1504	0	0,75
<i>bid-ask</i>	0,5092	0,4652	0	21,795
<i>coupon</i>	4,0391	1,7839	0,458	17,3
<i>log(amount issued)</i>	21,2165	1,802	14,5943	42,7135
<i>age</i>	1103,459	1101,996	0	9052
<i>age^2</i>	2432016	5684399	0	8,19E+07
<i>old</i>	0,7518	0,432	0	1
<i>investment</i>	0,9512	0,2155	0	1
<i>imp bid-ask</i>	0,2546	0,2326	0	10,8975
<i>gamma</i>	0,1328	0,152	0,0025	2,0468
<i>termfactor</i>	1,3345	0,6821	-0,2208	3,37
<i>deffactor</i>	0,0001	0,0056	-0,0484	0,0494

Table 3: Describing the characteristics of our variables of interest

Model 1 estimated with our liquidity proxies using random effects and robust standard errors									
Variables	1	2	3	4	5	6	7	8	9
<i>cdsindex</i>	0,2861 (21,48)***	0,2947 (18,43)***	0,2857 (20,49)***	0,2823 (19,91)***	0,2836 (21,29)***	0,3098 (18,27)***	0,249 (17,89)***	0,2861 (21,48)***	0,2861 (21,48)***
<i>vol index</i>	0,4175 (2,73)***	0,5097 (2,63)***	0,3927 (2,55)**	0,4049 (2,62)***	0,3719 (2,44)**	0,5092 (2,68)***	0,4106 (2,83)***	0,4174 (2,73)***	0,4173 (2,73)***
<i>duration</i>	2,8777 (2,29)**	3,1797 (2,01)**	2,6626 (2,79)***	2,5247 (2,66)***	2,3634 (1,89)*	4,7115 (1,81)*	2,5296 (2,10)**	2,8815 (2,29)**	2,8752 (2,29)**
<i>credit rating</i>	-26,1977 (-11,52)***	-21,7924 (-12,31)***	-26,1941 (-11,50)***	-26,0904 (-11,48)***	-26,1242 (-11,51)***	-28,971 (-11,64)***	-24,8761 (-11,60)***	-25,9361 (-11,41)***	-21,3412 (-12,19)***
γ		382,8972 (2,68)***							
<i>age</i>			-0,0016 (-0,30)	-0,0113 (-1,77)*					
<i>age²</i>				3.36e-06 (2,14)**					
<i>old</i>					-6,7319 (-2,65)***				
<i>log(amount issued)</i>						2,4697 (2,02)**			
<i>bid-ask</i>							39,5549 (3,97)***		
<i>imp bid-ask</i>									
<i>% zero</i>								74,6765 (3,12)***	
<i>coupon</i>									21,6832 (7,11)***
R ² total	0,4183	0,57	0,4189	0,4135	0,4191	0,4299	0,4474	0,4192	0,4514

Table 4: Regression output for ‘Model I’.

When examining the regression output we can see that there is one measure that stands out from the rest and that is γ , which is included in (2) in table (4). If we look at the separate regressions we can see that in all cases the *cdsindex* is significant at the one percent level, which makes the variable econometrically significant³.

The coefficient for this variable varies between 0.249 – 0.3098 between the 10 different versions of ‘Model I’ which is to be interpreted as if the *cdsindex* increases by one unit the spread between the corporate bond and its benchmark increases by 0.249 – 0.3098 basis points. As seen in table (3) the standard deviation from its’ mean is 94.31 which would imply that the effect on the yield spread from a change in *cdsindex* of one standard deviation is between 23.48–29.22 bps. The simple reasoning is that when the market credit risk increases, so does the average yield spread which is quite an intuitive result.

Volindex has a positive sign over all regressions and is always significant at the five percent level. The economic intuition is that when the volatility in the market increases, the market risk increases and investors will want a premium for the increased risk. Thus the yield spread will increase. The coefficients span between 0.3719 and 0.5097 and a change in volatility by

³ For the following reasoning of the regression output, unless something else is stated is being conducted on a ceteris paribus basis.

one standard deviation would increase the yield spread by 2.54 to 3.48 bps. Compared to many of our other variables this effect is not of great economical importance, implying that the market volatility hardly affects investors' assessment of corporate bonds. Another reason for the result might be that market risk is already captured by either *cdsindex* or any of the bond specific characteristics.

The first augmentation in our model is the duration. The coefficient for the duration spans between 2.3643 – 4.7115. The mean for the duration in our sample is 4.8938 with a standard deviation of 3.00, these statistics suggest that this variable lacks economic significance when explaining the spread from the benchmark. Statistically it is still significant in all of our regressions at least at the ten percent level. The coefficient for the variable is consistent with our ex-ante expectations meaning that a higher duration leads to a higher spread.

The second bond-specific characteristic that we chose to include is the credit rating. This variable seems to be of great importance in our results. As can be seen in table (3) the coefficient lies between -28.917 – -21.3412. This shows us that the credit rating is not only highly significant in a statistical sense but also in an economic sense, especially when considering the properties for credit rating. If we compare a AAA-rated bond to a CCC-, which is the highest and lowest in our sample, the difference in credit rating would account for an astonishing difference in spread between 384.1bps to 520.5 bps. Most of our sample consists of investment-grade bonds so the difference in credit rating is typically not that large, however, the credit rating, which is a proxy for the default risk, is a major part of the explanation of the spread.

The first of our proxies for illiquidity is γ . This proxy has both a statistical and economic significance. The standard deviation and mean for γ is 0.1520 and 0.1328 respectively meaning that one standard deviation change in γ would change the spread by 58.20 bps. If we bear this in mind when looking at figure (1) we can see that an impact of 58.20 bps has a large economic significance as well. These results indicate that the liquidity has quite a large effect on the yield spread.

Our second liquidity measure, *age*, lacks both economic and statistic significance but when we include *age*², we achieve statistical significance. The regression output tells us that the age of a bond is convex in its relationship to the yield spread. When examining the mean age of our sample, *age* and *age*² would decrease the yield spread by 4.3bps therefore we conclude that it more or less lacks economic significance.

A related measure is the binary variable *old*, which is defined as 1 for all bonds older than 2 years, which is true for roughly 75% of all observations in our sample. An old bond has on average a negative effect on the yield spread of 6.73. It is estimated to be statistically significant but the economic significance is rather small.

The logged amount is our next illiquidity measure, the coefficient for this measure is 2.47 and it exhibits statistical significance at the five percent level. The effect of a standard deviation from the mean is a change of 4.45 bps, which might be regarded as of low economic significance.

Our fifth measure is the reported bid-ask spread. This measure has an economic significance as well as being statistically significant at the one percent level. When the bid-ask spread increases by one standard deviation, the spread increases by 18.6 bps.

The sixth illiquidity measure is the percentage of non-trading days. Using the, by now familiar, mean/standard deviation analysis we can see that a standard deviation change increases the yield spread by 11.2 bps. It is statistically significant at the one percent level and could possibly have a large effect on the spread for the very infrequently traded bonds. The intuition is quite clear and should come as no surprise; a bond that rarely trades could potentially be hard to liquidate. For an investor to hold this type of bond she would require a premium.

Our last proxy of illiquidity is the size of the coupon. As mentioned earlier our ex-ante expectations of this measure was two folded but the regression result is quite clear. From the output we can conclude that a higher coupon is linked to a higher yield spread. The coefficient is both statistically and economically significant, a one standard deviation from its mean would generate a premium of 38.7 bps.

Also notable from the regression output of 'Model I' are the reported goodness of fit. Most of the versions yield similar R^2 except for the one including γ . All of the measures examined add little explanatory power to the basic model excluding γ , which increases the overall goodness of fit by 15% compared to the basic version of 'Model I'.

Results from testing the additional explanatory power of adding another proxie to model 1								
	Y	age	age+age^2	old	log(amount issued)	bid-ask	% zero	coupon
Stand alone	382,8972 (2,68)***	-0,0016 (-0,30)	-0,0113 (-1,77)*	-6,7319 (-2,65)***	2,4697 (2,02)**	39,5549 (3,97)***	74,6765 (3,12)***	21,6832 (7,11)***
Y	-	-	-	-	-	-	-	-
age	Y***	-	-	-	-	-	-	-
age+age^2	Y***	-	-	-	-	-	-	-
old	Y*	age***	X	-	-	-	-	-
log(amount issued)	log(amount)*	X	X	old**	-	-	-	-
bid-ask	bid-ask*	bid-ask***	bid-ask**	*	*	-	-	-
% zero	Y***	% zero***	% zero***	*	log(amount)**	*	-	-
coupon	coupon*	coupon***	coupon***	*	coupon***	*	*	-
Rank	3	1	0	1	1	2	2	3

Table 5: Rank-Test of the illiquidity measures

In the above ranking of the illiquidity measures we have performed pairwise regressions, other than for age and age square where there are 3 variables since we found age to have a quadratic impact on the yield spread. The top and the left row consists of the different measures and in the intersection between two variables there are either “-“, “variable”, “X” or “*”. The “-“ sign is simply telling us that the combination already has been done, “X” denotes a regression where both measures are insignificant at the 10% level, “variable” shows which variable that is dominant and the number of “*”s’ following tell us the significance level of the additional variable. If there is no variable but only a “*” then both variables are significant at the one percent level and we cannot determine that any variable subsumes the other. On the last row is the assigned rank, the number of times the variable subsumed another measure at the five or ten percent significance level. The coefficients for each variable from a stand-alone regression together with their associated z-values are also reported.

Based on the described methodology we can conclude that two variables seems to dominate the others more often, these measures are γ and *coupon*. The coefficient for *coupon* is positive in all of our regression for ‘Model I’, which would imply that when the coupon increases then the spread against the benchmark increases as well. Since we already in the regression control for duration as well as credit rating our ex-ante expectations were that the effect of the coupon size would be relatively small. This is based on the fact that a higher coupon would decrease the time until the investor would recoup her investment and at the same time a higher coupon would imply that the investment bears a higher risk and want to be compensated for that. These are two effects that contradict each other and our ex-ante were then that they would cancel out each other, perhaps with a small lean towards a positive coefficient.

Following the same reasoning γ is also of both statistical and economical significance. Considering that there are mostly investment grade bonds in the sample this effect is rather large. The reason why γ seems to be an appropriate measure of illiquidity can be twofold. Either the properties of illiquidity are truly captured by the negative autocovariance or there might be some correlated noise in u_t .

Model 1 estimated for only speculative grade bonds only									
Variables	1	2	3	4	5	6	7	8	9
<i>cds index</i>	0,9709 (8,76)***	0,9989 (8,73)***	0,9249 (8,22)***	0,9303 (8,41)***	0,9528 (8,82)***	0,9709 (8,76)***	0,7248 (5,72)***	0,9709 (8,76)***	0,9709 (8,76)***
<i>vol index</i>	0,0891 (0,04)	-0,3223 (-0,12)	2,1568 (0,95)	2,1667 (0,96)	0,0519 (0,02)	0,09 (0,04)	0,7774 (0,38)	0,0893 (0,04)	0,0892 (0,04)
<i>duration</i>	-31,3383 (-1,33)	-30,6224 (-1,04)	23,9115 (1,01)	26,829 (1,14)	-37,868 (-1,38)	-31,3295 (-1,33)	-13,0105 (-0,67)	-31,3436 (-1,33)	-31,3363 (-1,33)
<i>credit rating</i>	-127,7421 (-6,45)***	-129,2819 (-7,71)***	-180,0604 (-5,79)***	-180,639 (-5,90)***	-123,878 (-6,21)***	-141,1654 (-6,54)***	-125,201 (-6,73)***	-127,369 (-6,59)	-124,201 (-4,34)***
γ		441,2206 (2,64)***							
<i>age</i>			0,178 (2,31)**	0,2246 (2,63)***					
<i>age^2</i>				-0,00001 (-0,78)					
<i>old</i>					-41,0345 (-1,06)				
<i>log(amount issued)</i>						-90,8761 (-3,01)***			
<i>bid-ask</i>							106,0629 (5,03)***		
<i>% zero</i>								-25,0456 (-0,2)	
<i>coupon</i>									3,6864 (0,20)
R^2 total	0,5485	0,6775	0,3956	0,3999	0,5347	0,5801	0,6016	0,5489	0,5482

Table 6: Regression output for ‘Model I’ conducted on speculative grade bonds”

Model 1 estimated for only investment grade bonds only									
Variables	1	2	3	4	5	6	7	8	9
<i>cds index</i>	0,2321 (31,86)***	0,2347 (26,71)***	0,2278 (31,52)***	0,2238 (30,60)***	0,2298 (31,60)***	0,24 (27,71)***	0,2283 (32,45)***	0,2321 (31,86)***	0,2321 (31,86)***
<i>vol index</i>	0,5453 (8,09)***	0,6603 (8,17)***	0,319 (6,53)***	0,3333 (6,76)***	0,5007 (7,72)***	0,5824 (7,60)***	0,4879 (7,67)***	0,5442 (8,08)***	0,5442 (8,08)***
<i>duration</i>	4,5 (6,16)***	4,7861 (5,08)***	2,6654 (3,92)***	2,5028 (3,7)***	4,051 (5,53)***	8,9453 (8,44)***	4,1183 (5,82)***	4,5017 (6,16)***	4,4962 (6,17)***
<i>credit rating</i>	-8,9744 (-21,83)	-8,6565 (-17,49)***	-8,866 (-21,10)***	-8,7223 (-21,01)***	-8,878 (-21,82)***	-9,6704 (-19,23)***	-8,8878 (-21,73)***	-9,042 (-22,25)***	-8,273 (-20,34)***
γ		47,9819 (2,22)**							
<i>age</i>			-0,014 (-6,12)***	-0,0257 (-7,83)***					
<i>age^2</i>				3,98E-06 (4,08)***					
<i>d/d</i>					-6,1813 (-3,85)***				
<i>log(amount issued)</i>						0,6582 (0,95)			
<i>bid-ask</i>							10,9163 (4,41)***		
<i>% zero</i>								18,567 (1,96)**	
<i>coupon</i>									6,6442 (8,61)***
R^2 total	0,395	0,4087	0,35	0,3801	0,3986	0,3807	0,3972	0,3935	0,4167

Table 7: Regression output for ‘Model I’ conducted on investment grade bonds

The main differences between the speculative-grade bonds and the investment-grade bonds are the magnitudes of the coefficients for *cdsindex*, *creditrating*, γ and *bid-ask*. The coefficients for these variables are inflated for the speculative-grade bonds meaning that a unit change in these would have a higher impact on the yield spread compared to the investment-grade bonds. This would imply that bonds with lower credit rating are more sensitive to credit risk and liquidity risk. Since the spreads for the speculative-grade bonds are higher than for the investment-grade bonds we would expect that the coefficients would be higher.

Also worth noting is that the duration, number of non-trading days and the size of the coupon becomes statistically insignificant in the regression where we only include speculative-grade bonds. Since this is a smaller sample than the ones with higher credit ratings this could be a factor in explaining these findings.

We can also see that the goodness of fit has increased for most of our regressions when only including the bonds with lower credit ratings compared to the composite regressions and that it is higher than for all of the regression with only high rated bonds. Even though that one should not solely rely on the goodness of fit, this could suggest that our chosen variables better explain the spread from the benchmark for speculative-grade bonds.

Model 2 estimated with our liquidity proxies using fixed effects and robust standard errors									
	1	2	3	4	5	6	7	8	9
<i>default factor</i>	27,8923 (2,09)*	43,8929 (2,41)**	24,5567 (1,84)*	24,0553 (1,81)*	27,0254 (2,02)**	32,6823 (1,67)*	33,2651 (2,55)**	25,9916 (1,95)*	22,3519 (1,68)*
<i>term factor</i>	19,7594 (106,40)***	18,2227 (71,14)***	19,832 (106,96)***	19,9539 (107,75)***	19,8629 (107,02)***	25,3641 (109,16)***	16,7687 (91,35)***	19,8179 (106,98)***	20,8212 (111,56)***
<i>duration</i>	4,3887 (54,07)***	3,6154 (32,27)***	3,7551 (45,63)***	3,6018 (43,78)***	4,2559 (52,39)***	8,1121 (45,27)***	3,5039 (43,80)***	4,3935 (54,30)***	4,633 (57,45)***
2011	72,3556 (49,71)***	85,8397 (37,27)***	83,9714 (56,87)***	82,4378 (55,89)***	71,378 (49,06)***	126,9233 (24,03)***	48,3008 (33,51)***	73,2751 (49,33)***	98,1537 (47,65)***
2012	206,7251 (144,87)***	206,8247 (103,76)***	229,8554 (151,50)***	226,0247 (148,87)***	205,5855 (144,12)***	421,2022 (82,70)***	175,9543 (124,07)***	191,6449 (131,77)***	293,0524 (145,21)***
2013	40,9407 (28,77)***	62,7532 (31,14)***	74,8149 (46,37)***	70,3908 (43,58)***	42,056 (29,56)***	337,8689 (66,67)***	18,1711 (12,89)***	36,7471 (25,28)***	145,0045 (72,04)***
2014	-32,8695 (-14,96)***	-1,8998 (-0,59)	6,5595 (2,77)***	1,9259 (0,81)	-30,189 (-13,74)***	283,0194 (38,99)***	-47,6428 (-22,11)***	-32,7819 (-14,59)***	95,1379 (31,46)***
Credit rating 2010						1,4443 (10,21)***			
credit rating 2011	-3,5556 (-44,67)***	-4,2081 (-37,81)***	-3,5867 (-45,13)***	-3,5525 (-44,76)***	-3,4418 (-43,22)***	-2,9683 (-22,5)***	-2,4703 (-31,36)***	-3,6172 (-45,47)***	-4,0672 (-47,21)***
credit rating 2012	-11,0512 (-143,25)***	-10,8317 (-100,39)***	-11,0852 (-143,93)***	-10,9362 (-142,01)***	-10,8725 (-140,63)***	-11,8073 (94,94)***	-9,6541 (-126,08)***	-10,742 (139,41)***	-13,0496 (-155,21)***
credit rating 2013	-3,7998 (-49,33)***	-4,6633 (-2,83)***	-3,8176 (-45,13)***	-3,615 (-46,97)***	-3,6541 (-47,38)***	-4,2735 (-35,70)***	-2,5771 (-33,75)***	-3,752 (-48,72)***	-6,2933 (-75,10)***
credit rating 2014	-0,2626 (-2,21)**	-1,5576 (-8,96)***	-0,261 (-2,20)**	-0,0358 (-0,30)	-0,1655 (-1,39)		0,6963 (5,97)***	-0,3184 (-2,67)***	-3,4448 (-26,35)***
Y		17,6713 (55,04)***							
age			-0,0327 (-44,32)***	-0,0432 (-55,21)***					
age^2				3,86E-06 (39,96)***					
old					-7,348 (-29,48)***				
og(amount issued) 2011						-2,0048 (-8,55)***			
og(amount issued) 2012						-8,5611 (-37,77)***			
og(amount issued) 2013						-12,3267 (-54,54)***			
og(amount issued) 2014						-13,5452 (-42,03)***			
bid-ask							41,7947 (174,21)***		
% zero 2011								8,2267 (3,45)***	
% zero 2012								118,8152 (48,76)***	
% zero 2013								43,4512 (18,38)***	
% zero 2014								17,2904	
coupon 2010									(5,17)***
coupon 2011									16,7591 (60,97)***
coupon 2012									13,5062 (53,83)***
coupon 2013									5,2067 (22,11)***
coupon 2014									3,1253 (14,01)***
R^2 total	0,2372	0,2588	0,1242	0,1475	0,2357	0,1673	0,274	0,2417	0,3339

Table 8: Regression output for ‘Model II’

For our version of the Fama and French two-factor model with bond-specific characteristics we see that most of our included variables show statistical significance. In broad strokes the coefficients for most variables are similar in signs and magnitudes to ‘Model I’. We will focus on the biggest differences but first we must point out the goodness of fit. It seems as if ‘Model

II' only captures about a fourth of the variation in our dependent variable and we will thus assign greater trust to 'Model I'.

The major difference between the two models is the magnitude of γ , which is significantly smaller in 'Model II' than in 'Model I'. Since many of our liquidity measures are time-invariant we chose to include them as interaction terms as mentioned in the methodology section. This makes us unable to do a direct comparison with many of our illiquidity measures. The other illiquidity measures as well as the time-variant factors are similar to 'Model I'.

Coefficient sign outcome for 'Model I'		
Variables	Expected sign	Outcome
γ	+	+
Bid-Ask spread	+	+
% zero/Missing prices	+	+
Age	+	*
Amount issued	-	+
Coupon	+	+

Table 9: Table depicting a summary of our expectations for the proxies introduced above

Table (9) is an extension to table (2) in which you will find our expectations for the illiquidity proxies and the outcomes for them from 'Model I'. As can be seen above most of the measures behaved as predicted. *Age* came out ambiguous and the measure is perhaps not the most intuitive or accurate proxy for liquidity, which is bound to have richer properties and be more complex than what the information of a bonds' age can offer.

More notable is perhaps that the coefficient for *amount issued* displayed the opposite sign from what we expected. Even though it was barely economically and statistically significant we were somewhat puzzled to find that larger issues are related to higher yield spreads. As our main purpose was not to establish the relationship between the size of a bond issuance and the yield spread we did not examine this finding any further.

5. Discussion and Conclusions

In this section we will draw conclusions based on the material presented above. We will also discuss possible development for further research.

Evaluating our results we can conclude that there have been liquidity premiums in the Scandinavian corporate bond market. We found both economical and econometrical significance in several of the proxies used and we believe that there might be some truth to each of them.

Following our definition of liquidity we believe that a perfect measurement should include a time component and perhaps a value component. In γ and in the autocovariance of the price changes of the bonds we capture, to some extent, a little bit of both. The time component is the persistence of the γ measure over time and the value component is reflected by the level of γ .

Following the reasoning of Bao, Pan & Wang when deriving γ one obvious objection to the measure is that the negative autocovariance will capture more than illiquidity effects. In fact their reasoning is based on the assumption that any deviation from the fundamental value stems from illiquidity which the authors of this paper finds highly implausible. As seen in our analysis, the yield spread from benchmark is to a high extent caused by credit concerns and other factors. Nevertheless, when controlling for both credit and interest rate risks we still find some explanatory power of γ giving at least some creditability to the measure.

When comparing the impact from the illiquidity measure, γ , in our study to that of Bao, Pan & Wang (2011) we find similar results. Their study showed that the effect of a standard deviation from the mean of γ changes the yield spread by 65 bps. In our study we find that the corresponding effect is 58.2 bps. The difference in the effect on the yield spread could arise from a number of factors. When looking at our sample, we have two recent crises, the large financial crisis and the more recent euro-crisis. Another factor could be the transparency of the market. In the US, the authors and presumably investors have access to TRACE, where all of the US corporate bond trades are registered. In our sample we have had to rely on the trades being reported by the institutions to Bloomberg. They also had access to intraday data, which were not available to us.

Another distinction with this study is that they achieve a higher goodness of fit. Their study shows that the measure for illiquidity accounts for most of the variation the dependant ranging

from 47%-60%. When they include their CDS-index they increase their R^2 between 13%-30%. These figures for goodness of fit are quite higher than the ones that we found. The reasons for this might stem from a better approximation of γ since they had intraday data. Another contradiction to that study is that we find significance in several others of our illiquidity proxies.

The market in the US is more transparent we are therefore somewhat puzzled about our findings that γ in the Nordic markets are lower than in the US. Since the markets in Scandinavia are less transparent we expected that there would be larger liquidity premiums due to information asymmetry.

When comparing to Hoeweling et al. it is noticeable that they find age to be the measure with highest liquidity premium whereas we find the relationship between age and yield spread to be convex and of low economical significance.

The extension of the Fama-French model seems to be unsuitable to use for the Nordic corporate bond market. This might be due to the lack of proper indices for the construction of the default factor. The previous study done by Houweling et. al. was conducted through the use of pooled OLS while our data was better fitted by using the fixed effects model which made the interpretation of the results noisier.

Even though that we do not find the goodness of fit that Bao, Pan and Wang finds, we still achieve a model that accounts for almost 60% of the variation in the yield spread. When we examine 'Model I' we can see that most of the illiquidity measures fails to add much in terms of R^2 except for γ . This combined with the rank-test leads to the conclusion that γ is the most suitable of our proxies for illiquidity.

Unfortunately we are not certain that γ suffices to fully explain the characteristics of liquidity, there might be other frictional and/or behavioural components that can better be captured by other variables. Perhaps any of our other measures is closer to the truth.

5.1.1 Suggestions for further research

Our main headache has been the lack of more detailed transaction data. Collecting such is not an easy task but we believe that it would greatly improve the understanding of the microstructure dynamics of the Nordic corporate bond markets. Many of the more recent studies done in this area tend to use high-frequency data and thus increasing the precision of estimates.

Due to the lack of better data we had to exclude several potentially important measures for instance Amihuds liquidity measure and the LOT-model which both require volumes.

Another suggestion would be to use swap curves instead of government bonds as benchmarks in the light of the Houweling et. al. argumentation that investors are switching to interest rate swaps.

Works Cited

- Amihud, Y., & Mendelson, H. (1991). Liquidity, Maturity, and the Yields on U.S. Treasury Securities. *The Journal of Finance* , 1411-1425.
- Bank of International Settlements. (2014, April 4). *Long-Term Rating Scales Comparison*. Retrieved April 4, 2014, from Bank of International Settlements Home Page: <http://www.bis.org/bcbs/qis/rating.htm>
- Bao, J., Pan, J., & Wang, J. (2011). The Illiquidity of Corporate Bonds. *The Journal of Finance* , 66 (3), 911-946.
- Brennan, M., & Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics* , 441-464.
- Breusch, T., & Pagan, A. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *Review of Economic Studies* , 239-253.
- Chen, N.-F., Roll, R., & Ross, S. (1986). Economic Forces and the Stock Market. *The Journal of Business* , 59 (3), 383-403.
- Drukker, D. (2003). Testing for Serial Correlation in linear panel-data models. *The Stata Journal* , 168-177.
- Durbin, J., & Watson, G. (1950). Testing for Serial Correlation in Least Squares Regression, I. *Biometrika* , 409-428.
- Fama, E., & French, K. (1993). Common Risk Factors in the returns on Stocks and Bonds. *Journal of Financial Economics* (33), 3-56.
- Fisher, L. (1959). Determinants of risk premiums on corporate bonds. *Journal of Political Economy* , 529-546.
- Gebhardt, W. R., Hvidkjaer, S., & Swaminathan, B. (2005). The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics* , 85-114.
- Goldreich, D., Hanke, B., & Nath, P. (2005). The Price of Future Liquidity: Time-Varying Liquidity in the U.S. Treasury Market. *Review of Finance* , 1-32.
- Gopalan, R., Kadan, O., & Pevzner, M. (2012). Asset Liquidity and Stock Liquidity. *Journal of Finance and quantitative analysis* , 333-364.
- Grossman, S., & Miller, M. (1988). Liquidity and Market Structure. *The Journal of Finance* , 617-633.
- Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica* , 46, 1251-1271.
- Hillier, D., Ross, S., Westerfield, R., Jaffe, J., & Bradford, J. (2010). *Corporate Finance* (First European Edition ed.). Maidenhead: McGraw-Hill Education.

- Hoechle. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal* , 281-312.
- Houweling, P., Mentink, A., & Vorst, T. (2003). *How to Measure Corporate Bond Liquidity*. Amsterdam: Tinbergen Institute.
- Huang, J., & Wang, J. (2009). Liquidity and Market Crashes. *Review of Financial Studies* , 2607-2643.
- Hull, J. C. (2012). *Options, futures, and other derivatives*. Harlow: Pearson Education Ltd.
- Kyle, A. (1985). Continuous Auctions and Insider Trading. *Econometrica* , 1315 - 1336.
- McCulloch, J. (1975). An Estimate of the Liquidity Premium . *Journal of Political Economy* , 95-119.
- Newbold, P., Carlson, W., & Thorne, B. (2010). *Statistics for Business and Economics: Global Edition* (Vol. 7). Upper Saddle River, New Jersey: Pearson Education.
- Niederhoffer, V., & Osborne, M. (1966). Market Making and Reversal on the Stock Exchange. *Journal of the American Statistical Association* , 897-916.
- Palyi, M. (1936). *Liquidity*. Minnesota: Minnesota Bankers Association.
- Pastor, L., & Stambaugh, R. (2003). Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* , 642-685.
- Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance* , 1127-1139.
- Sarig, O., & Warga, A. (1989). Some Empirical Estimates of the Risk Structure of Interest Rate. *The Journal of Finance* , 1351-1360.
- Wiggins, V., & Poi, B. (2013, June). *Stata / FAQ: Testing for heteroskedasticity and autocorrelation*. Retrieved April 23, 2014, from Stata homepage: <http://www.stata.com/support/faqs/statistics/panel-level-heteroskedasticity-and-autocorrelation/>
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.
- Wooldridge, J. (2009). *Introductory Econometrics: A Modern Approach* (Vol. 4). Mason, Ohio: South-Western Cengage Learning.

Appendix 1

AAA	21	
AA+	20	
AA	19	
AA-	18	
A+	17	Investment-grade
A	16	bonds
A-	15	
BBB+	14	
BBB	13	
BBB-	12	
BB+	11	
BB	10	
BB-	9	
B+	8	
B	7	Speculative-grade
B-	6	bonds
CCC+	5	
CCC	4	
CCC-	3	
CC	2	
C	1	

Table A1: Conversion of the credit rating to numerical values and grouping

Appendix 2

Econometric Tests for 'Model I'

LM-test for Heteroskedasticity	
LR chi2(1230)	1414047
Prob > chi2	0.000
A Test for Heteroskedasticity for 'Model I' where: H_0 = Homoskedasticity	
Test for Serial Correlation	
F (1 , 1220)	85.304
Prob > F	0.000
A Test for Autocorrelation for 'Model I' where: H_0 = No Autocorrelation	
Hausman Test	
Chi2 (3)	3.940
Prob > Chibar	0.2680
A Test for Efficiency of Random Effects for 'Model I' where H_0 : Random Effects not most Efficient	
LM-test for Random Effects	
Chibar (01)	7.2e+07
Prob > Chibar	0.000
A Test for Random Effects for 'Model I' where: H_0 = No Random Effects	
DW-test after cluster	
DW-Statistic	2.072793
A Test for Autocorrelation for 'Model I' where: $DW(4, 1862) > 1.91217$ = No Autocorrelation	

Appendix 3

Econometric Test Results for Model 'II'

LM-test for Heteroskedasticity	
LR chi2(1230)	1269663.60
Prob > chi2	0.000
A Test for Heteroskedasticity for 'Model II' where: H_0 = Homoskedasticity	
Test for Serial Correlation	
F (1 , 1220)	57.705
Prob > F	0.000
A Test for Autocorrelation for 'Model II' where: H_0 = No Autocorrelation	
Hausman Test	
Chi2 (3)	119.54
Prob > Chibar	0.000
A Test for Efficiency of Random Effects for 'Model II' where H_0 : Random Effects not most Efficient	
DW-test after cluster	
DW-Statistic	2.668159
A Test for Autocorrelation for 'Model II' where: $DW(10, 1862) > 1.91964$ = No Autocorrelation	