UiT

NORGES
ARKTISKE
UNIVERSITET

Handelshøgskolen i Tromsø

Can stock returns be used as a proxy for corporate *bond yields?*

—

Stian Ek Grønning
Masteroppgave i Samfunnsøkonomi… Mai 2014

Contents

Introduction

This paper investigates the relationship between stocks and bonds issued by the same firm. Knowing a predictable relationship between two markets will be interesting for any investor, especially if this relationship is explained by a time factor. Suppose we could estimate the price of a derivate in terms of its underlying asset and this relationship was delayed in time. It would then be possible to forecast the future prices in of one of the markets, and hence, there could be arbitrage opportunities.

I expect to find some correlation between stocks and bonds issued by the same firm because they are both functions of the firm's risk. Suppose a firm experiences a period of adverse business and runs with a considerable deficit. The risk of bankruptcy will increase and hence it is less desirable to own the company. The stock will then be traded at a lower price. The same applies for the firms bonds. If the firm goes bankrupt, it will fail to make the promised payments to the bondholders. An investor will require a risk premium to hold a risky bond, and this is achieved only by buying it for a lower price. Thus, an increase in the overall risk will reduce the value of the bond.

By the same argument, I expect to find a correlation between the creditworthiness of the firm and the degree of correlation between the firm's bond yield and its share price. Rating agencies provide ratings describing the creditworthiness of corporate bonds. In this paper I have used Standard & Poor's rating system. AAA is the best grade. Following that comes AA, A, BBB, BB, B, CCC, CC C and D. The grade D means that the firm has failed to pay its debt and is bankrupt. Only 1.1 % of the top rated firms have defaulted after 20 years, while over 80% of the C-rated firms have defaulted in the same period. This results in the first null hypothesis:

H_0: There is no correlation between a firm's rating and the degree of correlation between its change in yield and stock return

I also expect that if there is a substantial amount of information asymmetry among the investors, the informed could choose one particular market at the expense of the other. It is then possible that there may be a time lagged relationship between the markets. Hence, the second null hypothesis:

H_0: *There are no time lags between the stock and the bond movement within the same firm*

Further, the reminder of this paper consists of six sections. The second is a literature review where I discuss what others have found. The third section contains some important background theory. This is included to help the reader to understand the theoretical concepts in this paper. In the fourth section I discuss some theoretical concepts regarding the relationship between the yield and the stocks within the same firm. The fifth section, methodology and results, is a statistical analysis where I investigate the relationship between stocks and bonds using various regression analyses. Here I present the models used for estimation, give some background theory on the procedures and present my results. In the sixth section, discussion, I will present some ideas regarding the analysis. The seventh section concludes the paper.

Existing literature

Kwan (1995), the basis for this study, have studied this relationship and found several significant correlations. He finds that changes in bond yields are significantly (at the 0.1% level) and negatively correlated with the issuing firm's contemporaneous and lagged stock returns, but there are no correlation between the bond yield and the leading stock return. This means that when new information is made available it impacts the stock market first. If the market were frictionless, it would be expected that the prices of the bonds and stocks would adjust to this change instantly and thus simultaneously.

Variable	Q1	Q2	Q3	Q4
R_t	-0.3235	-0.242	-0.2944	-0.1212
R_{t-1}	-0.137	-0.2434	-0.2811	-0.2533
R^2	0.28	0.34	0.42	0.47

The regression coefficients of the correlation between stocks and bonds. Contemporaneous and lagged based on firm size

He also finds that the firm size has an explanatory value. He investigates 327 firm's weekly returns and yields and splits them into

four quartiles according to size. The R^2 increases with the firm size, from 0.28 in the quartile for the smallest firms, to 0.47 for the largest firms. All values are significant at the 0.1% or 1% level. The coefficients don't change much in magnitude over the sizes, but the correlation has an increasing explanatory value.

He further finds that there are a significant correlation between the credit rating of the firms and their correlation between bonds and stocks. The yield of top rated bond is uncorrelated with the firm's stock price. The bond yield depends only on the risk free rate. In the table below the coefficients of his regression are presented (t-values in parentheses). The correlation increases when the rating goes down.

Variable	AAA	AA	A	BBB	BB	B
R_t	-0.1963	-0.0878	-0.1033	-0.3489	-0.5011	-0.4079
	(-1.47)	(-2.64)	(-3.58)	(-8.76)	(-7.24)	(-5.85)
R_{t-1}	-0.2015	-0.1981	-0.2483	-0.3313	-0.3309	-0.1656
	(-1.84)	(-6.37)	(-9.15)	(-5.08)	(-5.08)	(-5.55)
R^2	0.61	0.5	0.41	0.4	0.04	0.04

The regression coefficients of the correlation between stocks and bonds based on the rating of the firm

One issue here is the R^2 of BB- and B-rated firms. If we follow the scale downwards from AA, the R^2 diminishes slightly and suddenly drops to almost zero. The coefficients are still highly significant. This implies that there are other factors than the rating that explains the correlation. A possible explanation is that the rating covers aspects like the coverage-, leverage-, liquidity-, profitability- and cash flow-to-debt-ratio, not just the volatility of the firm.

Hotchkiss and Ronen: (2002) have analyzed the informational efficiency of the high yield corporate bond market using daily and hourly price data. They did not find support for the claim that stock portfolio returns lead bond portfolio returns, but they managed to detect a contemporaneous correlation between stock and bond returns.

Norden & Weber (2009) have done an even more comprehensive analysis than Kwan's. They find the same general results. In addition they find that American firms have a stronger negative correlation with their bonds than European firms, and that telecommunication firms exhibits a much stronger correlations than others.

Background Theory

Stocks

Ownership of a stock (sometimes called a share) represents a claim to one proportion a firm's assets and earnings. In other words, stockholders are the owner of the firm. If a firm has 10,000 stocks outstanding and an investor owns 5,000 of these, he owns half of the firm. In an efficient market, the combined value of the stocks will represent the firm's value. If a business reports of a negative change of its profits or an unexpected deficit, the business is considered less valuable and the price of the stocks will decline. Corporations have limited liability. This means that the owner of a stock can never be held personally responsibility for more than the value of the stock. This means that a creditor cannot demand that a stockholder puts up more money to pay of the firm's debt.

Options

Options have properties that are highly interesting in the pursuit of an answer to the research question. There are different options associated with the ownership of bonds and stocks. I will return to this issue later.

An option is a contract which gives the owner the right to buy or sell an underlying asset for a specific price before a specific time. The issuer of the contract is obligated to buy or sell (depending on the contract) the underlying asset if the owner chooses to use this right. We then say that the option is exercised.

There are two kinds of options. A call option gives the owner the right to buy the underlying and a put option gives the right to sell. Usually, one contract is an agreement to buy or sell 100 stocks. Unlike futures and forwards, where the owner is obligated to exercise the contract, the option only gives the owner the right, not the obligation to exercise it. This freedom may be of great value for the owner, and hence the contract has intrinsic value. The underlying asset may be several things, such as stocks, currencies, stock indices, futures etc.

Valuing options at the* exercise *date

A call option will have value on the final day of its lifetime if the underlying asset is worth more than the exercise price. If the value of the stock is 140, and the strike price is 100, the

owner can then use the option to buy the stock for 100 and sell it immediately for 140 (assuming no commission) and then earn 40 in no time. Hence, the option is worth 40, assuming there are no commission fees. A put option works in the opposite way. If the stock is worth 100 and the strike price is 140, the owner sells the stock for 140 and buys it back again for 100, and earns 40. This will also mean that the option is valueless if the value of the underlying asset is on the "wrong" side of the strike price. If the stock is worth 130 and the strike price is 140, it would make no sense to exercise a call option.

An option cannot take any value. A call can never be worth more than the underlying asset. If this was not true, you could earn money by buying the stock and selling the option. A put can never be worth more than the strike price.

Valuing options before the exercise date

The determination of the value of an option before the expiration date is a much more complicated matter. This problem was solved by Fisher Black and Myron Scholes in 1973 in their famous work "The Pricing of Options and Corporate Liabilities". In addition to the current value of the underlying asset (S_0) and the strike price (K), the value now depends on some more factors:

- **T**: Time to expiration of the option
- **r** : The risk free rate
- **σ**: The volatility of the underlying asset

The value of the option on the expiration date and an earlier point in time may differ due to the possibility of a change in the value of the underlying asset. For example, if the strike price on a call option is 80, the underlying is worth 70 and there are three months remaining of the contract, there is always a possibility for a positive change in the underlying's price of more than 10 during the remaining time. An option is therefore never totally worthless. There will always be at least a faint hope for a profitable price change. The amount of time to expiration and the volatility of the value of the underlying are hence important factors in estimating the possibility of a desirable price development. The price itself is calculated with Black & Scholes differential equation. I will not go through this process in detail here.

Bonds

A bond is an instrument of indebtedness of the bond issuer to the holder, a way of acquiring capital. The issuer can be a government, municipality or corporation. It basically works like this: The bond is written and sold for a certain value to a holder. At the end of the bonds lifetime the issuer pays the holder back with interest. The ownership of bonds can be transferred in the secondary market.

There are several types of bonds. The most common is the coupon bond. Here is the issuer obligated to make fixed payments to the holder according to an agreed upon indenture in addition to the principal (face value) in the end of the bonds lifetime. Another type is the discount bond. This bond is sold at a lower price than the principal. The holders return is the difference between the price and the face value. There are also other types of bonds which coupon payments are not fixed, but are derivatives of other financial sizes, like various indices and rates.

Some bonds are callable. This means that the issuer can choose to pay off the principal early. If the macro environment changes and the general cost of money decline, the bond will be relatively costlier for the issuer. The bond may then be called, and new bonds with different indentures are written. This implies a risk for the holder of the bond because he might not find an equally good investment opportunity if the bond is called, and hence the required return is higher if the bond is callable.

Stocks and bonds are both securities. One important difference between the two is that the owners of the stocks are investors, and the owners of the bonds are lenders. The bondholders have priority and will be repaid before the stockholders in the event of bankruptcy. We will return to this matter later on.

Valuing bonds

The price of a safe bond is the present value of all the cash flow generated by the issuer to the holder. The discount factor is here the risk free rate since the bond is risk free. In practice the risk free rate may vary over different periods, and this needs to be incorporated into the calculation.

$$\text{Bond value} = \sum_{t=1}^{T} \frac{\text{Coupon}}{(1+r)^t} + \frac{\text{Face value}}{(1+r)^T}$$

The price of the bond changes over time because the holder of the bond earns continuously a share of the next coupon payment. This is called the accrued interest. If a coupon bond is sold one second before a coupon payment is made, the owner of the bond in the recent period is entitled to the incurred rates. The invoice price equals the stated price plus the accrued interest.

Yield to Maturity

The most commonly used measure for the return of the bond is called yield to maturity. This is the internal rate of return (IRR) on the investment in the bond, assuming that all coupons are reinvested in that yield.

The yield curve

The yield curve shows the relationship between the time to maturity and the yield. The most common shape of this curve is depicted below. The concavity of the curve shows that bonds with a longer maturity promise a higher yield than those with a shorter maturity. This is because investors require a risk premium due to the uncertainty of future events. The further the maturity is into the future, the higher risk premium is demanded. The shape of the yield curve changes over time and is of great interest because it gives an idea of the expected future interest rates and economic activity. Therefore, the slope of the yield curve is important. If the slope is steep, there is a large gap between short- and long-term bonds.
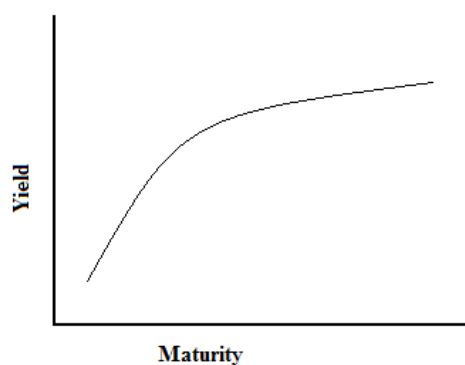
Theoretical analysis

Valuing bonds when there is a significant risk of default using option pricing theory:

One essential aspect regarding bonds is the difference between the promised and actual payment. If the payment is risk free, such as Norwegian government bonds, the promised and actual payments are always the same. Due to the general risk of default, corporate bonds have on average a higher promised yield than government bonds. The compensation for this risk, the risk premium, will depend on the particular corporation issuing the bond, the current market conditions, the rate of safe bonds, etc.

Merton (November 1973) presents a theory of how to value bonds when there is a significant probability of default. He defines the risk as "…the possible gains of losses to bondholders as a result of (unanticipated) changes in the probability of default…". He further claims that a difference in price between two bonds with equal term structures will solely be caused by the difference in the probability of default. Both Merton (spring 1973) and Black and Scholes (1973) recognize that this can be done with the same method as with option pricing.

They have 8 assumptions for their theory to work. They are as follows:

1. There is no transaction cost or taxes.
2. There are a sufficient number of investors participating in the market who can buy or sell as much as they wish of an asset to the market price.
3. The borrowing and the lending rate are the same.
4. Short-sales of all assets are possible.
5. Trading takes place continuously in time.
6. The Modigiliani-Miller theorem that the value of the firm is invariant to its capital structure obtains.
7. The term structure is flat with certainty. This means that future values are discounted continuously
8. The dynamics of the value of the firm, V, can be described as

$$dV = (\alpha V - C)dt + \sigma V dz$$

where α is the expected rate of return on the firm per unit of time, C is the total dollar payouts by the firm to its shareholders, σ^2 is the variance of the return on the firm and dz is a standard

Gauss-Wiener process. That all these assumptions hold in every point in time may perhaps be a little too much to wish for, but the author claims that assumption 1-4 may be substantially weakened and argues that the rest of them will hold.

The theory is described by assuming there exists a security whose market value, Y, can be written as a function of the value of the firm and time. The dynamics of the security's value can be written as:

$$1) \quad dY = (\alpha_y Y - C_y)dt + \sigma_y V dz_y$$

Since $Y = F(V, t)$, there is a relationship between α_y, σ_y and dz_y and the corresponding variables described in assumption number 8. By use of Itô's lemma, we can write the dynamics of Y as:

$$2) \quad dY = F_v dV + \frac{1}{2} F_{vv}(dV)^2 + F_t$$
$$= \left(\frac{1}{2}\sigma^2 V^2 F_{vv} + (\alpha V - C)F_v + F_t\right)dt + \sigma V F_v dz$$

Subscripts denote the partial derivatives. If we compare I. and II: we can show that the instantaneous returns on Y and V are perfectly correlated:

$$3) \quad \alpha_y Y = \alpha_y F \equiv \frac{1}{2}\sigma^2 V^2 F_{vv} + (\alpha V - C)F_v + F_t + C_y$$
$$\alpha_y Y = \alpha_y F \equiv \sigma V F_v$$
$$dz_y \equiv dz$$

Now, assume that we take a loan and invest the money in the firm and the particular security. w_1 is the amount of money invested in the firm, w_2 is the amount invested in the security and w_3 is the size of the loan. ($w_3 = w_1 + w_2$). If dx is the instantaneous return from the portfolio, we have:

$$4) \quad dx = w_1 \frac{(dV + Cdt)}{V} + w_2 \frac{(dY + C_y)}{Y} + w_3 r dt$$

$$= \left(w_1(\alpha - r) + w_2(\alpha_y - r)\right)dt + w_1\sigma dz + w_2\sigma_y dz_y$$

$$= \left(w_1(\alpha - r) + w_2(\alpha_y - r)\right)dt + \left(w_1\sigma + w_2\sigma_y\right)dz$$

The investment strategy is chosen so that the stochastic parts of the firm and the security cancel each other. If this is accomplished the portfolio is nonstochastic. If there are no arbitrage profits, the expected return on a zero net investment is zero.

5)

    a. $w_1\sigma + w_2\sigma_y = 0$

    b. $w_1(\alpha - r) + w_2(\alpha - r) = 0$

If $w_j \neq 0$, the solution to equations in 5 exists if and only if

6) $\left(\frac{\alpha - r}{\sigma}\right) = \left(\frac{\alpha_y - r}{\sigma_y}\right)$

Now, substitute for α_y and σ_y from equation 3 and rewrite:

$$\frac{\alpha - r}{\sigma} = \frac{\frac{1}{2}\sigma^2 V^2 \, F_{vv} + (\alpha V - C)F_v + F_t + C_y - rF}{\sigma V F_v}$$

Rewrite again and simplify:

7) $0 = \frac{1}{2}\sigma^2 V^2 \, F_{vv} + (rV - C)F_v + F_t + C_y - rF$

This is a parabolic partial differential equation for F. It must be satisfied by all securities whose values can be written as a function of the value of the firm and time. To solve this equation, we need an initial condition and two boundary conditions. These conditions are what distinguish one security from another (e.g., the debt of a firm from its equity). Equation 7 says that in addition to time and the value of the firm, the function for the value of the security Y=F(•), depends on the risk free rate, the variance of the firms value, the payout policy of the firm and the promised payout policy to the holders of the security. It also shows

that F does not depend on α, the expected return of the firm, the utility functions of the investors and other assets that are not present in the portfolio.

I will now review the simplest case of theoretical bond pricing using what we have found so far. Suppose there is a firm with only two classes of claims: A simple discount bond and equity. If F is the value of the debt issue, we can write equation 7 as:

$$8) \quad 0 = \frac{1}{2}\sigma^2 V^2 F_{vv} + rVF_v - rF - F_\tau$$

Where $C_y = 0$ because this as a zero coupon bond and $C = 0$ if we assume that the firm cannot issue any new claims on the firm, pay cash dividends or do share repurchase prior the time T. $\tau \equiv T - t$, the length of time until maturity, hence $F_t = -F_\tau$. We will need two boundary conditions and an initial condition to solve equation 8. By definition, $V \equiv F(V, \tau) + f(V, \tau)$ where f is the equity's value. F and f cannot take negative values due to the limited liability, so we have the lower boundary:

$$F(0, \tau) = f(0, \tau) = 0$$

Moreover, the bond cannot be worth more than the issuing firm. Hence the upper boundary:

$$F(V, \tau)/V \leq 1$$

The initial condition arises from the following argument: The firm promises to pay an amount of money to the bondholders at the maturity date T. If the firm defaults and the payment is not met, the bondholder takes over the company and the shareholders lose everything. So, if we have that $V(T) > B$, where B is the amount of money to be paid to the bondholders at the time T, the firm will pay out, and the value of the equity will be $V(T) - B > 0$. If $V(T) < B$, the firm defaults and the bondholders effectively acquires the firm. Thus, the initial condition for the debt at $\tau = 0$ is:

$$F(V, 0) = \min[V, B]$$

Now the equation can be solved with the standard method of separation of variables. However, these calculations can be avoided by looking at a problem already solved by the literature. The value of equity $f(V, \tau)$ will in this case equal $V - F(V, \tau)$. If we substitute for F in equation 8 and use the same boundary conditions, we have that:

9) $\frac{1}{2}\sigma^2 V^2 f_{vv} + rVf_v - rf - f_\tau = 0$

Subject to the initial condition:

$$F(V, 0) = \max[0, V - B]$$

If we rearrange 9) and change the notation to $f = w$, $f_v = w_1$, $f_{vv} = w_{11}$, $v^2 = \sigma^2$, $V = x$ and $w_2 = -f_\tau$, we have exactly the same as equation 7 in (Black & Scholes 1973. P.643):

$$w_2 = \frac{1}{2}\sigma^2 V^2 f_{vv} + rVf_v - rf - f_\tau$$

This is an equation for a European call option on a non-dividend-paying stock where the firm value in equation 9 corresponds to stock price, and B corresponds to the exercise price. I have not taken on the challenge to do this calculation, but I have shown that this is a valid way of estimating the price of a high risk bond.

Estimating fictitious stock price for a non-publicly traded firm when the bond price is known:

As mentioned if the former section of this paper, when a firm issues bonds, the bondholders effectively acquires the firm and the stockholders obtain an option to buy it back. If the stockholders can provide more equity, they have the option to pay off the bondholders whenever they find it suitable. This is because most corporate bonds are callable. The stockholders have in effect bought a call option on the assets of the firm from the bondholders. Owning a corporate bond is, by the same argument, equivalent to holding a safe bond and at the same time giving the firm's stockholders a put option on the firm's assets.

Thus, the stockholders are given the opportunity to sell the assets to the bondholders if the firm defaults.

The value of the put is the value of limited liability for the stockholders. They cannot be held economically responsible for the firm's debt. So, if the firm defaults, the stockholders will walk away from their firm's debt and hand over the firm's assets to its creditors/bondholders. The stockholders lose their entire deposit of money and the bondholders lose their claim to the firm minus the value of the assets. This will imply that the call option is valuable if the business runs profitably and the firm is able to pay its debt. On the other hand, in the event of a default, the put option will be the valuable one. But, as previously mentioned, due to the fact that financial and economic conditions may change over time, none of the options will be completely worthless before the default is a fact.

If we now look at the put-call parity,

"Value of call + present value of exercise price = value of put + stock price"

 and regard the "present value of exercise price" as the present value the promised payment to bondholders for sure next year, the parity can be rewritten as:

"Value of call + present value of promised payment to bondholder = value of put + stock price"

Since holding a corporate bond is equivalent to holding a safe bond and at the same time giving the firm's stockholders a put option on the firm's assets, we can again rewrite the equation:

"Present value of promised payment to bondholder - value of put = stock price - value of call = value of the bond"

To sum up: The put option has a positive value, and the call option is worthless on the maturity date if the assets are worth less than the debt. If there is a positive amount of time before the debt expires, the value of the options is not necessarily the same as on the time of

maturity. This means that if the bond price is known in the exchange, and the stock is not publicly traded, a theoretical stock price can easily be calculated if the price of the option is known. This can be done with the Black & Scholes formula for option valuation. The only unobservable variable here is the volatility of the stock, but it is possible to do reasonable assumptions here.

Methodology and results

To investigate the relationship between the stock return and the bond yield issued by the same firm, I have done various regression analyses on stock and bond data. I have tried several different models, here including variables on level form, on difference form, with different time lags, distributed lag models, autoregressive distributed lag models and various regression using dummy variables.

Data description

The dataset comprises of data for 47 consecutive business days, containing the price of the stock and the yield of the bond for 27 American companies. Moreover, I have used the closing price of the stock and the yield calculated from the last corresponding bond transaction of the day. All stock prices are adjusted for dividends, merges and splits. It would be preferable to use intraday data, but due to the fact that bonds have a very low trade frequency, good intraday data does not exist for very few bonds, and seldom for several days in a row. The requirements for a firm to be included in the analysis are as follows:

1. Its stock has to be traded publicly in a stock exchange.

2. The firm's bond has to be traded approximately every day. If there were periods of no trading, the data would provide poor information and the two time series would not be suitable for comparison.

3. It must be rated by Standard & Poor's prior to the first day of estimation and the rating score could not change within the period. I could have used Moody's or Fitch's rating system as well. They are equal for the purpose of this paper. The choice of S&P is

arbitrary.

4. The bonds maturity date needs to be in 2017 or close to this year. It would be natural to use bonds with approximately the same maturity so the differences in yield among the firms would not be due to the yield curve. The choice of 2017 is arbitrary, but my choice fell on this year since it was a little more frequent represented, and hence it was easier to find bonds with a common maturity year.

5. The bonds are a coupon bond with fixed payments. The change in yield cannot be subject to special coupon calculation.

6. The bonds are of senior security level. This means that if a firm has issued several bonds with different security levels and faces payment difficulties, the payouts will happen after a priority scheme where the senior security bonds will be paid out before other subordinated bonds.

7. The bonds need to be callable. The callability will impose a risk for the holder, and hence it affects the price and the yield. This is a very common property of corporate bonds, and I assume that the presence of this feature may affect the change in yield, and hence I have used only callable bonds.

Beyond these requirements, the choice of firms is arbitrary. I have not taken parameters like firm size or age, amount of debt, industry sector or geographical location into account. This is a weakness of this paper, but I had to draw a line for the scope of this paper somewhere.

The reason why I have used the yield instead of the price of the bond has two main reasons. Firstly, the yield to maturity is the most common measure for the return on bonds, and I have no reason to deviate from other literature. Secondly, one price of the bond may give different yields, depending on the amount of time until maturity. Hence, an investor that requires a certain return on his portfolio will place different bids on the bond at different point in time. The table on the next page shows the yield arising from a purchase of a bond for the price in the left column at the time in the top row. The bond in the example matures on the 7th of April 2024, pays a 5% coupon and has a face value of 100. The time premium/ yield curve is not taken into consideration in this example, but the argumentation is still valid.

Price	4.7.2014	4.7.2015	4.7.2016	4.7.2017	4.7.2018	4.7.2019	4.7.2020	4.7.2021	4.7.2022	4.7.2023
90	6.37 %	6.48 %	6.63 %	6.82 %	7.07 %	7.43 %	7.97 %	8.87 %	10.69 %	16.23 %
100	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %	5.00 %
110	3.79 %	3.68 %	3.55 %	3.38 %	3.16 %	2.84 %	2.37 %	1.57 %	0.00 %	-4.65 %

Suppose that a firm has a constant risk, and the investor has a certain opinion about the future of this firm and then requires a yield of minimum 3% for this bond. The table shows that the investor would pay 110 for the bond in 2018, but not in 2019. Hence, the change in price between two subsequent periods is not 100% comparable to the change in price between two other subsequent periods if there is a considerable time span between them. This difference will be larger if the price is far from the face value. Although, this may me nit- picking for the sake of this paper, but it will be an important aspect for research on long maturities.

Regression analysis

The various relationships investigated in this paper are estimated with regression analyses. I will now explain how this works, address some common problems one might encounter, and suggest some ways of solving such problems.

We are often interested in knowing how changes in economic variables (e.g., price) will affect other economic variables (e.g., consumption). A regression analysis is an approach for modeling this relationship, described in mathematical terms. The most basic method, the linear regression, explains the dependent variable (y) as a function of the estimated parameters ($\beta_1, \beta_2 \dots$), the independent variable x_i, and an unexplained error term ε_i:

$$y_{i=} \beta_0 + \beta_i x_i + \varepsilon_i$$

Given a sample for the population we wish to investigate, we obtain an estimate of the parameters:

$$\hat{y} = \widehat{\beta_0} + \widehat{\beta_i} x_i$$

The residual, $\varepsilon_i = y_i - \hat{y}_i$ is the difference between the predicted and the observed dependent variable. The ordinary least square method (OLS) obtains the parameter values that

give the smallest sum of squared errors, and hence, the most accurate line through the plotted data.

There are several assumptions that are necessary for the regression result to be unbiased and consistent. They are:

- The errors are random, follow a normal distribution and have a mean of zero.
- The errors are uncorrelated.
- The variance of the errors is constant through time.
- The sample is representative for the population.
- The independent variables are independent of each other. It should not be possible to express one variable as a linear combination another.
- The independent variable must take at least two different values.

There are several other aspects that need to be considered before the analysis is done. The input data should be diagnosed to reveal possible problems (we shall return to this issue shortly), and one should check if some form of data correction is needed. A thorough analysis of the dataset will also help to determine which regression model that should be applied. It is further important to distinguish between cross-section data, which is data on a number of economic units at a particular point in time, and time-series data, data collected over time on one particular economic unit. This paper predominantly investigates time-series data. The first issue one should then consider is autocorrelation and the possible presence of a unit root.

Autocorrelation
Autocorrelation is a problem that arises in time series regression when the dependent variable is a function of its former value. It can be described as:

$$y_{t =} \beta_0 + \beta_1 y_{t-1} + \varepsilon_i$$

This is a violation of the assumption of constant variance in the errors. This is detected because it produces autocorrelation in the observable residuals. This means that the current error affects not just the current value of the dependent variable, but also its future values. As an example, suppose that a natural disaster creates a fear of shortage of a certain resource,

driving the price of this resource up for an extended range of time. This event is not predicted by the model, and hence it is a part of the error term.

Autocorrelation in the error can be due to an autocorrelated dependent variable whose autocorrelation is not sufficiently explained by the independent variables and their lags. It can also be caused by omission of an autocorrelated independent variable.

Autocorrelation does not make the coefficients (betas) biased, but, when the autocorrelation is positive, the standard errors tends to be underestimated, and hence the t-value tends to be overestimated. This may lead us to a conclusion where a coefficient is significant when it is actually not and hence committing a type II error. If it is possible to correctly model the autocorrelated errors, an alternative estimator with a lower variance may exist. A lower variance gives a higher probability to obtain a more accurate coefficient estimate.

How do we detect and measure autocorrelation? The most common procedure is to calculate the Durbin- Watson d- statistic. This is a bound test for the null hypothesis that the errors are serially uncorrelated. The alternative is that the errors follow a first order autoregressive process. The test statistic is:

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=2}^{T} e_t^2}$$

T is the number of observations and e_t is the residual error at time t. The statistic ranges from 0 to 4, where low values indicate that successive error terms are, on average, close in value to one another. They are then said to be positively correlated. High values indicate the opposite. The error terms are then, on average, much different in value from one another, and they are negatively correlated. As mentioned earlier, the errors need to be normally distributed with a zero mean for the regression analysis to be valid. A rule of thumb is that the value of the statistic should lie between 1 and 3, preferably 2. The exact critical values depend on the number of observations, the desired level of significance and the number of independent variables. The values can be found in appendices of statistical texts. Two values are reported, and if the statistic lies between them, the autocorrelation test is inconclusive. If the null hypothesis of no autocorrelation is rejected, the whole regression analysis may be useless because it does not give a trustworthy result. If we use a lagged dependent variable as an

independent variable, the d- statistic is not always reliable. Durbin's h-statistic should then be used. This is an asymptotic normally distributed statistic for large samples. This means that is follows a standard normal distribution. The regression can then be tested with the null hypothesis of no autocorrelation against the 2-sided alternative of autocorrelated errors. Hence, at a 5% level, the decision rule is if `-1.96 < h < 1.96` do not reject the null hypothesis. This statistic cannot always be computed because the square root of a negative number may be required

So what do we do when we reject the null hypothesis of no autocorrelation? Donald Cochrane and Guy Orcutt published a seminal paper in 1949 which describes a procedure to adjust a linear model for serial correlation in the error term. The steps in the estimation procedure are as follows:

Step 1:
Use the ordinary least squares- regression to obtain the residuals

Step 2:
Run the regression: $e_t = \rho e_{t-1} + v_t$

This gives the least squares estimate of ρ as $\hat{\rho} = \frac{\sum_{t=2}^{N} e_t e_{t-1}}{\sum_{t=2}^{N} e_{t-1}^2}$

Step 3:
Use the estimate $\hat{\rho}$ to obtain observations of Y* and X* as:

$Y_1^* = \sqrt{1 - \hat{\rho}^2}Y_1$, $X_1^* = \sqrt{1 - \hat{\rho}^2}X_1$ and
$Y_t^* = Y_t - \hat{\rho}Y_{t-1}$, $X_t^* = X_t - \hat{\rho}X_{t-1}$, for t=2,3,...,N

An estimate of β is obtained from an OLS regression of Y* on X*. A new set of residuals is then calculated and steps 2 and 3 are repeated until successive estimates of ρ differ by less than 0.001. After the procedure is done, the Durbin- Watson statistic will be closer to 2. If the initial residuals were subject to a large extent of autocorrelation (D.W. close to 0 or 4), the correction may not be perfect, but the null hypothesis of zero autocorrelation will at least be inconclusive. If the amount of autocorrelation is very large, and the independent variables

seems not to be explaining the variation in the dependent variable, the variable could be nonstationary.

Stationary vs. nonstationary variables

A time series variable is stationary if its mean and variance is constant over time. This will imply that the covariance between two arbitrary values in the time series depends only on the time span between them.

Most economic variables are random. This means that it is not possible to perfectly predict it, so the true value is not known until it is observed. A model that produces such a time-series variable is called a stochastic process. A univariate time-series model is an example of a stochastic process where the value of a single variable is only dependent of its former values and past error terms. This can be described as:

$$y_t = \rho y_{t-1} + \upsilon_t \qquad |\rho| < 1$$

This is an autoregressive model where the errors are independent, with zero mean and with a constant variance. The fact that $|\rho| < 1$ implies that the process is stationary. The variance of this process can be shown to be the constant $\sigma_\upsilon^2 / (1 - \rho^2)$, which is not dependent of t.

Example of a random, stationary process with zero mean
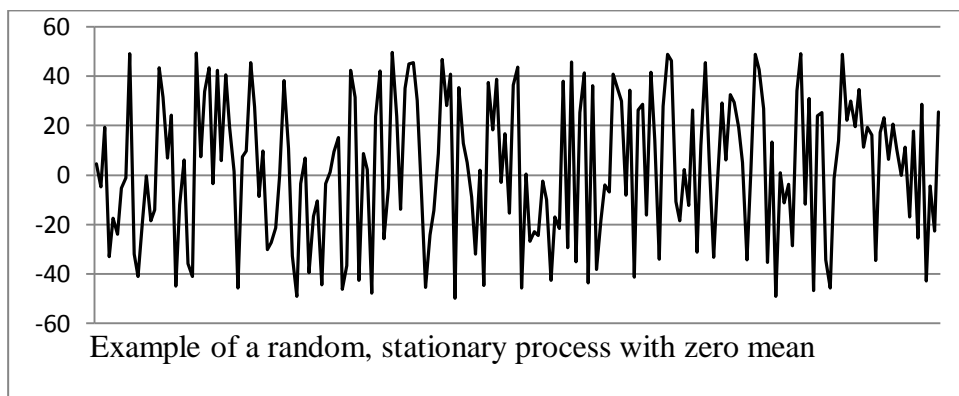
A non-stationary variable is described as not having the property of mean reversion. This means that it does not tend to return to its previous mean after a shock, like a stationary process will. Consider an autoregressive model fluctuating around a linear trend:

$$y_t = \alpha + \rho y_{t-1} + \lambda t + \upsilon_t$$

This process contain a growth term, λt, and its mean will depend on t. When $|\rho| < 1$, the process will be stationary around its trend.

Example of a nonstationary process with stationary trend
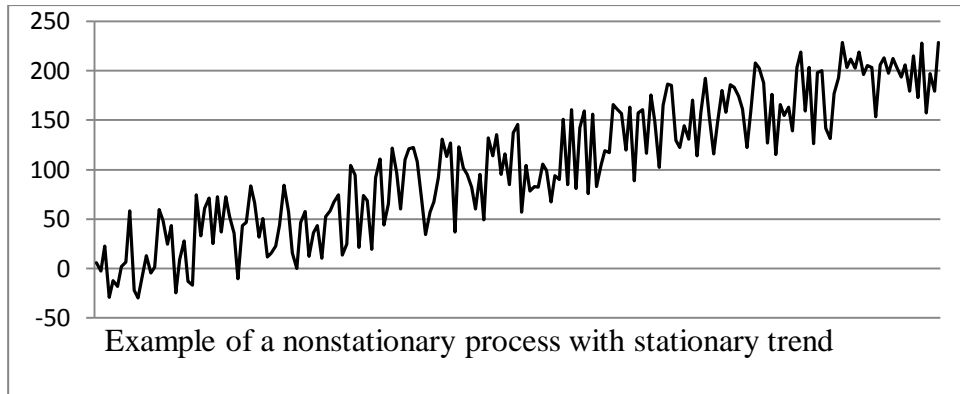
A process will also exhibit nonstationarity if $|\rho| = 1$. We then say the process has the property of a unit root. This means, if we disregard the intercept and the growth term for now, that the model takes the form:

$$y_t = y_{t-1} + \upsilon_t$$

This is called a random walk model. The value of y_t depends only on its former value plus the stochastic error term. This means that process evolves through time, and hence the variance and the mean will also follow the same randomness. Since the random error is added for each time, and this is the only reason for change in the variable besides it former value, the time-series will move in unpredictable directions, and it will be impossible to predict the variables next value. The mean values of subsamples will be dependent on the samples period. A unit root will cause problems in statistical inference. A regression can be highly inaccurate or spurious if one or more of the variables exhibits this kind of different "behavior" at different points in time.

Example of random walk model
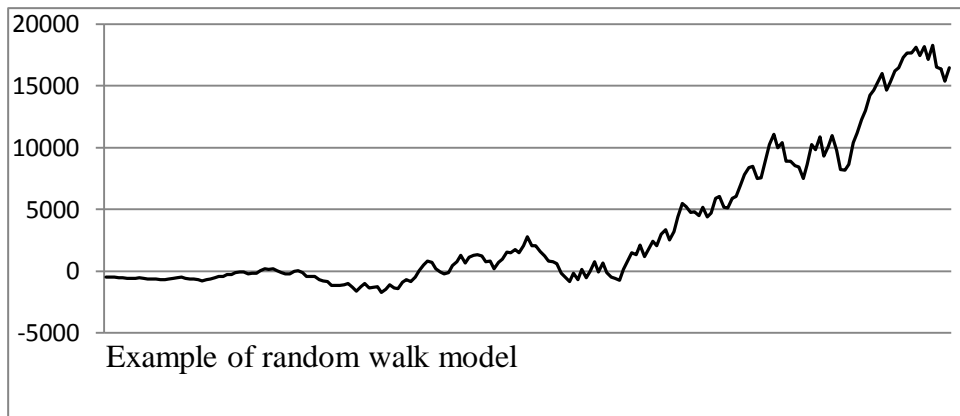
If nonstationary variables were to be used in a simple regression model, the results could indicate that there is a significant relationship between them when there actually is none. If the two time series have fairly similar shape, an OLS-regression could indicate a quite significant relationship, but in reality they just happen to drift in the same direction.

The Dickey-Fuller test

The tests for stationarity and unit root were developed by David Dickey and Wayne Fuller in 1979. As mentioned before, an autoregressive model like $y_t = \rho y_{t-1} + \upsilon_t$ is stationary when $|\rho| < 1$ and nonstationary when $|\rho| = 1$. This means that the interesting thing to examine here is the value of ρ. More accurately, we test if ρ is equal to one or significantly less than one.

$$H_0: \rho = 1 \text{ or } H_1: \rho < 1$$

Tests for this purpose are called unit root tests for stationarity. There are three different versions of the Dickey-Fuller test for a unit root. Note that the most common version of this test is on difference form:

1. $\Delta y_t = \beta_0 y_{t-1} + \varepsilon_i$
   This is a plain test for a unit root.
2. $\Delta y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_i$
   The second version is a test for a unit root with drift.
3. $\Delta y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 t + \varepsilon_i$
   The last one is a test for a unit root with drift and a time trend.

The procedure runs several tests to check the null hypotheses $H_0: \beta_0 = 0$, $H_0: \beta_0 = \beta_1 = 0$ and $H_0: \beta_0 = \beta_1 = \beta_2 = 0$. In unit root test on difference form we test of the coefficients equals zero, not 1. A problem with this test is that if the null hypothesis is true, y_t is nonstationary and has a variance that increases as the sample size increases. This will transform the distribution of the usual t-statistic. Therefore, we use another statistic called a τ (tau) statistic with its own unique critical values. We reject the null hypothesis of nonstationarity if $\tau \leq \tau_c$. In other words, in the case of $\tau \leq \tau_c$ the time series is stationary.

A further developed version of the test, called an augmented Dickey-Fuller test (ADF-test), allows for the possibility that the error term is autocorrelated. This problem may arise if the applied model fails to capture the full dynamic nature of the process by lacking some important lag terms. The number of sufficient lagged terms can be determined by examining the autocorrelation function of the residuals. In practice, the augmented Dickey-Fuller test is always used by statistical software to ensure that the error terms are uncorrelated.

Cointegration

An interesting feature of a model that exhibits a unit root, such as $y_t = y_{t-1} + \upsilon_t$, is that it becomes stationary if we take the first difference, which means that we add its former lag to the equation. It can then be written as $\Delta y_t = y_t - y_{t-1}$. Models that has this property is said to be integrated of order one, and is denoted as I(1). Stationary variables are integrated of order zero, denoted I(0). The order of integration is the number of times the time series has to be differentiated to make it stationary.

In general, nonstationary variables should not be used in a regression due to the possibility of obtaining a spurious result. However, there is an exception to this rule. If we have two variables which are nonstationary and integrated of order one, we expect their difference, or any linear combination of them, such as $e_t = y_t - \beta_1 - \beta_2 x_t$, to be integrated of order one as well. But there are exceptions where the linear combination is integrated of order zero. If this is the case, the said variables are cointegrated. This means that they share a common stochastic trend, and hence they never diverge too far from each other. If this is the case, the residuals must be stationary. The regression will then **not** be spurious. A Dickey-Fuller test, as previously described, can reveal the stationarity of the residuals.

To sum up:

$$H_0: \text{the series are not cointegrated} \iff \text{residuals are nonstationary}$$
$$H_1: \text{the series are cointegrated} \iff \text{residuals are stationary}$$

Multicollinearity

Multicollinearity is a phenomenon that arises in a multivariable regression analyses when two or more of the variables are highly correlated. This is a problem because it makes it difficult to make inferences about the individual coefficients. In a regression analysis we are often interested in knowing what happens with the dependent variable when we change one of the independent variables, and hold the other ones fixed. This will not give a sensible interpretation if there is a dynamic relationship between the independent variables. Here are some examples of issues that may arise if the model has a problem with highly correlated independent variables:

- An independent variables coefficient that is expected to be an important predictor may turn out to be insignificant or have the wrong sign.
- Omission or inclusion of an independent variable may change the values of the other regression coefficients drastically.
- The standard error of the affected variables may be estimated to be too large compared to if they were "alone" in the equation. If we test if the coefficient is equal to zero, we may be led to a failure to reject a false null hypothesis of no effect.

If the same pattern of multicollinearity is maintained through the time series, it may not be a severe problem for the overall model. The predictive power is not diminished as the dependent variable is a function of a bundle of independent variables. So we could chose to leave the model as is, despite multicollinearity, depending on the research question. If this is a problem, ridge regression or principal component regression can be used to solve the problem.

In time series data, the presence of multicollinearity between a variable and its lag will be the same as autocorrelation.

Time lagged variables

A lagged variable is a time series variable that has been moved in time. This is may be useful because the causality between cause and effect may take some time. If this is the case, and we look at the instantaneous relationship between x and y, we might not find the relationship between the variables. Therefore, we "push" one or several variable one or several steps in time, so that the regression is corrected for this "time error". To reveal a possible lagged relationship between the variables, it is beneficial to use some statistical software that simulates different lags and lags combinations, and then reports the best model based of information criterions.

A model that is relevant for this paper is the autoregressive distributed lag model: a model that contains both lagged x_t's and y_t's as independent variables. With *p* lags of *y* and *q* lags of *x*, the model can be written as.

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + v_t$$

Information Criterions

There could be a problem to determine the best model to use in econometrical analyses. Information criterions are measurements for the relative quality of a statistical model. They cannot be tested for since they don't provide a statistic in absolute sense. The criterions are based on a trade-off between goodness-of-fit and the complexity of the model. AIC is given by

$$AIC = \ln\left(\frac{SSE}{N}\right) + \frac{2K}{N}$$

and the Schwarz criterion is given by

$$SC = \ln\left(\frac{SSE}{N}\right) + \frac{K \ln N}{N}$$

Where:

- SSE: Sum of squared errors
- N: Number of observations
- K: Number of variables

The formulas are quite similar, but SC penalizes extra variables more heavily than AIC for N>8. One should use the model with the smallest value of AIC or SC. One important aspect regarding this method is that it will not tell if the model fits the data poorly in general, only that it fits the data better than others.

Burnham & Anderson (2002) recommend that AICc (AIC corrected for finite sample sizes) should be the preferred measurement for model selection regardless of sample size. This criterion is AIC with an addition:

$$\text{AICc} = \text{AIC} + \frac{2\text{K}(\text{K}+1)}{\text{N}-\text{K}-1}$$

AICc converges to AIC as N gets large. The use of AIC when N is not many times larger than k^2, will increase the probability of selecting models that have an excessive number of parameters, and this can be a problem in some cases.

Dummy variables

A dummy variable is binary. This means that it can only take two values, namely zero and one. It is used to present non-quantitative properties. This can be gender, model, location, color or a grade in an ordinal scale. The zero and the one indicates the presence or the absence of a property. Usually the variable D is defined as:

$$D = \begin{cases} 1 \text{ if the property is present} \\ 0 \text{ if the property is no present} \end{cases}$$

In a regression analysis, a dummy variable can be used to capture the changes in the intercept and/ or the slopes. Another neat property of the dummy variable is that it can be used to describe interactions between qualitative factors. Intercept variables for qualitative properties are additive. This means that the effect of each factor is summed up and added to the regression intercept. Hence, the regression model assumes that the dummy variables are independent of each other, but this is not always the case in reality. Though, if they are independent, it is easy to interpret the coefficients estimated from the regression. If the qualitative factor is present, just add its coefficient to the intercept.

If we are interested in knowing whether a qualitative factor is significant in the explanation of the variation of the dependent variable, we could perform an F-test to check if all the coefficients are equal to zero at the same time. If this is the case, the qualitative factor does not explain the variation of the dependent variable.

The regression models

This econometrical analysis is done in two main steps. The first is an analysis on each individual firm and the second part is a study of the relationship between the firms. I start with 10 different regression models and various correlation estimates of on each firm. I have done a kind of "model mining", where I have searched for a model that can explain the relationship between stocks and bonds issued by the same firm. This must not be mistaken for data mining, where one searches for data that fits a desired finding. Due to presence of autocorrelation, I have used the Cochrane- Orcutt iterative estimation in the first 8 models, but not it 9 and 10, where a lagged dependent variable is included in the regression. The neat property of C.O.- estimation method is that it corrects the errors for autocorrelation regardless of if this is a problem or not. This means that autocorrelation is fixed when it is present, and otherwise, the regression is the same as OLS- estimation.

I have used the following notations: Y is the yield, S is the price of the stock and ε is the error term. Since the literature suggests that the change in the yield is a function of the stock return, I have focused predominantly on this relationship. I have used the following regression models:

1. $Y = \beta S + \varepsilon$

   This is a simple regression on levels.

2. $\Delta Y = \beta \left(\frac{\Delta S}{S} \right) + \varepsilon$

   This regression is on the change in yield and the return from the stock as a percentage.

3. $\Delta Y = \beta_1 \left(\frac{\Delta S}{S} \right)_t + \beta_2 \left(\frac{\Delta S}{S} \right)_{t+1} + \varepsilon$

   Difference form. Contemporaneous and leading x-variable.

4. $\Delta Y = \beta_1 \left(\frac{\Delta S}{S} \right)_t + \beta_2 \left(\frac{\Delta S}{S} \right)_{t-1} + \varepsilon$

   Difference form. Contemporaneous and lagged x-variable.

5. $\Delta Y = \beta_1 \left(\frac{\Delta S}{S} \right)_{t+1} + \beta_2 \left(\frac{\Delta S}{S} \right)_t + \beta_3 \left(\frac{\Delta S}{S} \right)_{t-1} + \varepsilon$

   Difference form. Leading, contemporaneous and leading x-variable. I have chosen to do these three regressions (3., 4., and 5.) partly to investigate the multicollinearity present in the independent variables, and partly to check if some of the various lags capture the true timing property of the relationship between stocks and bonds.

6. $\Delta Y_{t-x} = \beta \left(\frac{\Delta S}{S} \right) + \varepsilon$

   Model number 6 is a series of regressions done in a loop. The first regression is the same as model 2. The second is the same but with a lagged dependent (dY) variable. The third loop is with the same variable lagged two times, and so on. This is done because I am trying to reveal the length a possible time lagged relationship between stocks and bonds. This method will show the most probable distance in if the relationships regardless of the significance of the variable. However, there are no economic theories that suggest that this time span will have a duration of several days, but I have checked out it nevertheless. The most ideal would be to do

the loop several times for each company where different number of time lags of the stock return (independent variable) would be included. I have not done this because this would require hundreds of regression, and it was not possible to find any software that could do this. It would also be convenient lag the variables both way, say for t+5 to t-5, but this was also not possible with the software, but I solved this problem with model 7.

I needed a method to determine which of the tested time lag(s) that would explain the time lag in the best manner. I solved this problem by using the Akaike information criterion (AIC) and the Schwarz criterion (SC).

7. $\Delta Y = \beta \left(\frac{\Delta S}{S} \right)_{t-x} + \varepsilon$ This model is principally the same as nr 6, but now is the independent variable $\left(\frac{\Delta S}{S} \right)_t$ lagged. This is the same as leading the dependent variable.

8. $\Delta Y = \beta \left(\frac{\Delta S}{S} \right)_{t-1} + \varepsilon$

The literature suggests that stock return leads the change in bond yield, and I have here tried to see if I could obtain the same result with one days lag.

9. $\Delta Y_t = \beta_1 \Delta Y_{t-1} + \beta_2 \left(\frac{\Delta S}{S} \right)_t + \varepsilon$

This is an autoregressive lag model on difference form. I have included this model to see if I could correct the autocorrelation by adding a lagged dependent variable to the equation. This is not important to answer to research question.

10. $Y_t = \beta_1 Y_{t-1} + \beta_2 \left(\frac{\Delta S}{S} \right)_t + \varepsilon$

This is the same as model 9, but on level form.

Results

First, all the variables were diagnosed with the augmented Dickey- Fuller test for unit root. As mentioned before, the presence of a unit root means that the variable exhibits a stochastic behavior, and hence the variance is not constant through time and one of the main preconditions for a regression analysis is violated. One third of the yield variables and almost all the stocks (as expected) had this property. This problem was overcome by doing the regressions mainly on difference form or checking the residuals for stationarity, and hence cointegration of the variables. But, since only a few firms tested positively for cointegration, I will focus on the findings on difference form. This is also the most common approach in the literature.

	BB+	BB	BB	B	B	CCC+	CCC	CCC	CCC-	CC
β	-2.081	-1.212	1.883	-1.041	-8.149	0.553	-6.536	-5.941	-3.687	-2.841
P	0.032	0.062	0.045	0.164	0.015	0.826	0.672	0.272	0.371	0.000
R^2	0.362	0.325	0.237	0.142	0.241	0.05	0.22	0.342	0.141	0.492

Table: Regression with the model $\Delta Y = \beta \left(\frac{\Delta S}{S} \right) + \varepsilon$

Model 2 found, as expected, that the highest rated firms show no significant correlations between the change in yield and the return of the stocks (on the 5% level). The highest rated firm with a significant relationship has the rating BB+, which is just below investment grade, and is the highest speculative grade. All the other firms that have this relationship also have a lower rating than BB+. With other words, the rating plays a role here. I have tested my results and managed to successfully reject my first null hypothesis. This was done with a dummy-variable regression to obtain the coefficients. Then I tested if all the coefficients were equal to zero at the same time. The reported p-value was 0.000. I could not number the various rating grades and use an ordinary OLS. This is because the grade scale is not numerical, but ordinal. The "distance" in-between the grades are not known.

I expected to find a negative coefficient in model 2. It is reasonable to believe that an increase in the stock price will be associated with an overall reduced risk for the company, and hence a lower risk for the bonds to default on the payments. In this case we would also expect the yield to decline due to a reduced risk premium. Although this is the case for most of the

companies, there is one exception where the coefficient is positive, and I cannot explain this and it may be a random finding.

Model 3, 4 and 5, the models with different combinations of leading, contemporaneous and lagged variables did not find any new significant coefficients. One interesting result in model 3 is that the inclusion of an insignificant time lag to the contemporaneous and significant variable changed the beta-coefficient. The reported p-value was changed considerably. The contemporaneous variable was made insignificant. This indicates that there is a problem with multicollinearity in these models, and hence I have not managed to find a model that captures the true relationship of the timing property.

Model number 6 and 7 estimated the length of the best lag between the variables, and found, besides the contemporaneous already found, that there were a significant lag for the stock return of one day for **one** of the firms. This may be due to chance. The next model looked specifically at the one-day lag. The same result was found here. In other words, I have not successfully managed to detect a lead-lag relationship between the change in yield and the stock return. This means that I have not successfully managed to reject my second null hypothesis of no time lags.

Model number 9 and 10 came up with some results that may be of interest. Highly significant lagged ΔY's as independent variables seems to be a common feature here. This would be expected on level form since the time series is autocorrelated, but this implies that some yield-time series are autocorrelated at a higher order.

I did not find the relationships for any rating to be significant on the 0.1% level, like Kwan (1995) and Norden & Weber (2009) did. Their accurate results may be due to a better dataset. I have constructed my dataset manually by looking at transaction data, and because this is an extremely time consuming work, I have a maximum of 4 firms per grade, including the sub grades with a minus or a plus. Other papers have used data from hundreds of firms over a longer time span, and this will provide richer information.

It is also possible that they have used better and more complex models. Both Norden & Weber (2009) and Hotchkiss & Ronen (2002) have used a vector autoregressive model. This

captures both lead and lag relationships within and between stationary variables in a simultaneous multivariate framework.

Norden & Weber (2009) have already found out that there are parameters other than the rating that explains the relationship between stocks and bonds. Firm size, leverage ratio and industry sector are some parameters that plays a role here.

Discussion

The time plot for ΔY for Beazer Homes has a fairly consistent variance. It is tested to be stationary. The mean of Y has a weak declining tendency, but is fairly stable, and has a value of 5.32% in the given period. It is also not correlated with the firm's stock return, which has a totally different trend behavior. The price of the bond must necessarily follow the same basic pattern as the yield. Hence the question: Why is this bond bought and sold for several different prices?
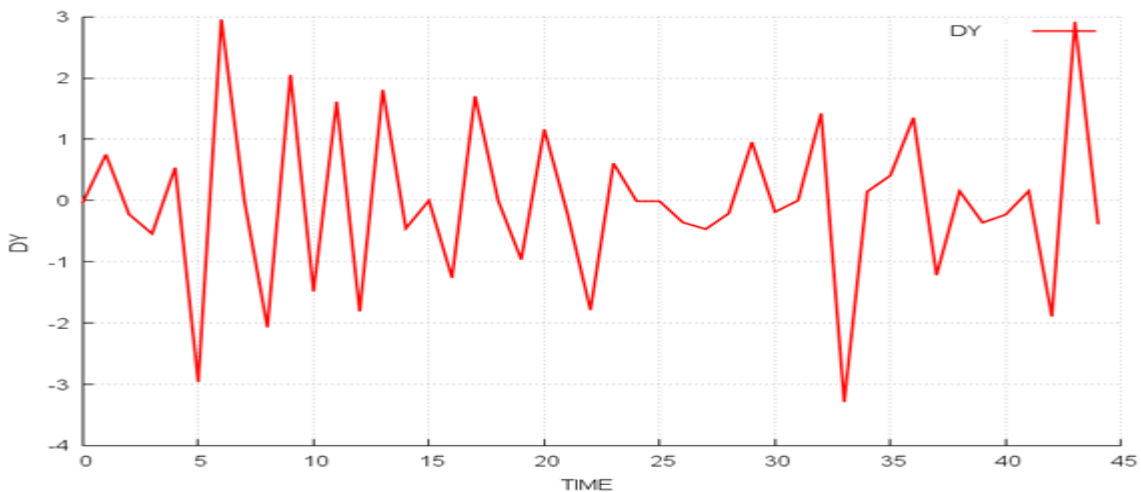
The change in yield for Beazer Homes

The correlation matrix (next page) for the different time lags shows that the subsequent ΔY's are negatively correlated to each other. This means that a change in the yield tends to be followed by a change with the opposite sign. I will therefore assume that the yield, and hence the price, will be possible to forecast to a certain extent. The direction of the price change should at least be possible to predict quite often. Remember that this pattern is consistent

throughout the whole period of more than two months. It is possible that this pattern has existed before and after the given period to. I will return to this subject in the discussion part.

LdY0	1.00			
LdY1	-0.57	1.00		
LdY2	0.08	-0.59	1.00	
LdY3	0.01	0.08	-0.56	1.00
	LdY0	LdY1	LdY2	LdY3

The correlation matrix for the time lags of the change in yield.

A visual inspection of the plot of the development of various yields and the corresponding stock prices shows that they often coincide to a large extent, although their relationship is not statistically significant. This may imply that there is an economic relationship between them, although they are not statistically significant. Due to different opinions and expectations among the investors, the stochastic part accounts for a large portion of the movement. I think a logic reason to why the investors disagree on the price of a bond is that the method for calculating a bond price under risk is a very difficult task. I will therefore assume that a considerable proportion of the bond bids are based of belief and guessing. Again, this is just an assumption.

An interesting question that would naturally follow the research question is what do the firms that exhibit a significant relationship between the stocks and the bonds have in common? The only common property I have successfully managed to detect is that these firms have one huge peak in the stock return. The magnitude of this movement in the level of the stock price is so large that it probably instantaneously affects the overall risk and hence the creditworthiness of the firm and therefore also the bonds. The plot for ΔY and stock return for the firm Best Buy (next page) illustrates this perfectly. We can clearly see that the points are almost perfectly collected around the origin in a circular fashion, but there is ONE point in the far North West corner. On January 15th the firm published a financial statement bearing negative news, and the stocks fell rapidly by 28.6% and the yield increased by 6%. If this one point is omitted from the regression, the result is no longer significant. This is also the case for several other firms.
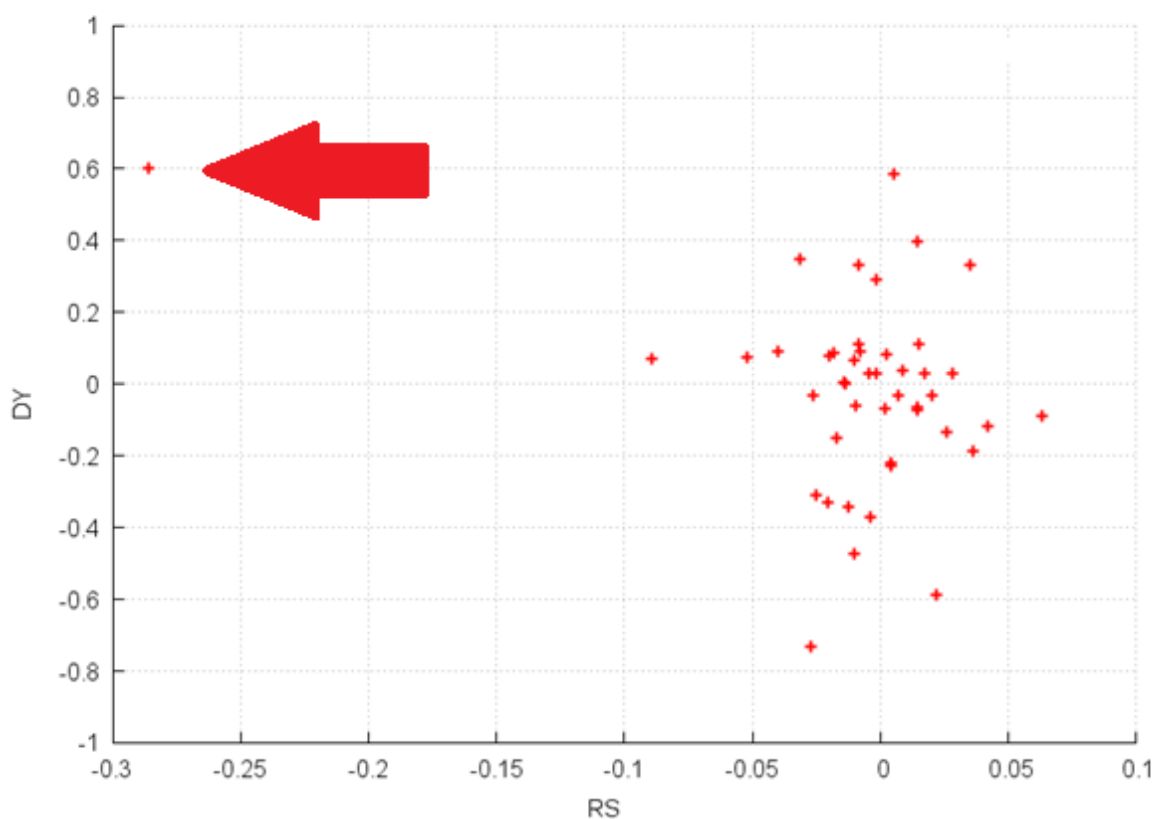
Plot of the stock return against the change in yield

Conclusion

In this paper, I have investigated the relationship between the change in yield and stock returns for 27 firms for 47 consecutive business days. Firstly, I have found that there is an inverse relationship between the rating of firms and the degree of correlation between the stock return and the change in yield for bonds issued by the same firm. Secondly, I have not successfully managed to demonstrate that this relationship has a timing property in general. It was found for only one of the firms. Thirdly, it seems that the most common feature among the firms that exhibits a significant relationship is a large and sudden change of the stock price, but I have not statistical evidence for this claim.

References

Articles:

Banz, R.W. (1981): *The Relationship Between Return and Market Value of Common Stocks*

Barsky, R. B. (1989): *Why Don't the Prices of Stocks and Bonds Move Together?*

Black, F. and Scholes, M. (1973): *The Pricing of Options and Corporate Liabilities.*

Burnham, K. P. and Anderson, D. R. (2002): *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach* (2nd ed.)

Cochrane, D. and Orcutt, G. H. (1949): *Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms*

Durbin, J. (1970): *Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables*"

Fama et al (1989): *Business Conditions and Expected Return on Stocks and Bonds*

Keim, S. and R. F. Stambaugh (1986): *Predicting Returns in the Stock and Bond Markets*

Kwan, S. H. (1995): *Firm-Specific Information and the Correlation Between Individual Stocks and Bonds.*

Merton, R.C (spring 1973): *A Rational Theory of Option Pricing*

Merton, R.C (November 1973): *On the pricing of corporate debt: The risk structure of interest rates.*

Norden, L. and Weber, M (2009): *The Co-movement of Credit Default Swap, Bond and Stock Markets: an Empirical Analysis.*

Books:

Hull, J.C (2012): *Options, futures, And other Derivatives. Eighth edition. Chapter 4, 9, 10 and 23*

Bodie, Kane, Marcus (2011): *Investments and Portfolio Management. Ninth edition. Chapter 14 and 15*

Myers, B. (2003): *Principles of corporate finance. Seventh edition. Chapter 1, 20, 22 and 24.*

Carter Hill, R., Griffiths, William E., Lim Guay, C. (2012): *Principles of Econometrics. Fourth Edition. Chapter 6, 7, 9 and Appendix 9A*

Lind, Marchal, Wathen (2010): Statistical Techniques in Business and Economics.

Web pages:
http://www.standardandpoors.com/ratings/definitions-and-faqs/en/us

Appendix1. The dataset:

Dato	JNJ	JNJ.GO	MSFT	MSFT3926357	AAPL	AAPL4001806	GOOG	GOOG.AB	PFE	PFE4013015	KO	KO3830397	ORCL	ORCL3919232	BA	BA.GBZ	CAT	CAT.HOQ	ADBE	ADBE.GA
02.jan	90.38	1.374	36.88	1.206	549.84	1.691	1113.12	0.549	30.21	1.024	40.34	1.546	37.72	1.399	135.9	0.189	89.29	2.362	59.29	0.602
03.jan	91.19	1.374	36.64	1.346	537.76	1.786	1105.00	0.356	30.27	1.007	40.14	1.546	37.62	1.373	136.85	0.438	89.24	2.374	59.16	0.602
06.jan	91.67	1.485	35.86	1.265	540.69	1.568	1117.32	0.713	30.3	1.014	39.95	1.702	37.47	1.446	137.63	0.557	88.06	2.32	58.12	0.602
07.jan	93.62	1.361	36.14	1.267	536.83	1.678	1138.86	0.713	30.49	1.049	40.07	1.677	37.85	1.378	139.72	0.725	88.35	2.309	58.97	0.81
08.jan	93.49	1.327	35.49	1.238	540.23	1.767	1141.23	0.611	30.69	0.895	39.63	1.782	37.72	1.476	140.03	0.058	88.56	2.39	58.9	0.722
09.jan	94.05	1.28	35.27	1.306	533.33	1.814	1130.24	0.595	30.67	1.03	39.42	1.786	37.65	1.59	141.33	0.474	89.13	2.287	59.09	0.611
10.jan	94.06	1.24	35.77	1.276	529.77	1.673	1130.18	0.595	30.44	0.898	39.82	1.543	38.11	1.388	141.1	0.942	89.92	2.206	59.53	0.715
13.jan	93.82	1.172	34.72	1.303	532.54	1.665	1122.98	0.538	30.29	0.977	39.22	1.487	37.75	1.43	139.91	0.446	89.31	2.198	58.6	0.758
14.jan	94.04	1.296	35.51	1.258	543.14	1.876	1149.40	0.581	30.74	1.017	39.38	1.254	38.21	1.454	139.22	0.462	89.97	2.263	60.37	1.121
15.jan	94.12	1.261	36.49	1.258	554.04	1.709	1148.62	0.491	30.92	1.042	39.45	1.551	38.41	1.472	139.83	0.483	91.81	2.27	61.68	0.466
16.jan	93.96	1.169	36.62	1.326	550.95	1.672	1156.22	0.59	30.91	0.848	39.4	1.595	38.29	1.45	139.42	0.45	92.01	2.249	61.63	0.71
17.jan	94.38	0.963	36.11	1.222	537.45	1.653	1150.53	0.672	30.83	1.004	38.97	1.53	38.21	1.45	139.67	0.515	91.44	2.243	61.37	0.71
21.jan	93.36	1.14	35.9	1.15	545.8	1.658	1163.70	0.124	30.97	1.085	39.61	1.839	38.11	1.108	140.88	0.651	90.6	2.257	60.85	0.639
22.jan	93.65	1.198	35.66	1.268	548.23	1.787	1165.02	0.62	31,01	1.03	39.59	1.636	37.98	1.507	143.56	0.448	89.64	2.31	61.77	0.676
23.jan	92.08	1.141	35.79	1.255	552.87	1.702	1160.10	0.473	30.75	0.99	38.93	1.629	38.15	1.443	140.52	0.468	88.48	2.231	60.88	0.648
24.jan	89.96	1.356	36.54	1.262	542.82	1.634	1123.83	0.626	29.84	1.022	38.53	1.532	37.11	1.517	135.88	0.668	86.17	2.216	59.09	0.657
27.jan	89.3	1.095	35.76	1.484	547.22	1.674	1101.23	0.642	29.41	1.062	38.43	1.533	36.49	1.465	136.59	0.382	91.29	2.232	58.36	0.651
28.jan	89.46	1.134	36	0.945	503.49	1.659	1123.01	0.627	30.17	1.023	38.56	1.574	37.1	1.44	136.32	0.382	92.47	2.2	59.11	0.572
29.jan	88.26	1.147	36.39	1.285	497.77	1.613	1106.92	0.617	29.86	1.023	37.6	1.408	36.97	1.41	129.05	0.474	90.62	2.101	58.46	0.533
30.jan	88.86	1.103	36.59	1.191	496.81	1.622	1135.39	0.602	30.57	0.924	37.87	1.475	37.4	1.41	125.82	0.396	93.2	2.151	59.39	0.565
31.jan	87.84	1.015	37.56	1.118	497.62	1.565	1180.97	0.602	30.15	0.974	37.52	1.404	36.9	1.39	124.56	0.939	93.91	2.128	59.19	0.966
03.feb	86.16	1.076	36.21	1.195	498.55	1.531	1133.43	0.517	30.35	0.57	36.91	1.449	35.84	1.328	122.39	0.359	92.42	2.059	58.09	0.781
04.feb	86	1.147	36.08	0.837	505.76	1.459	1138.16	0.588	31.18	0.845	37.19	1.376	35.96	1.314	121.36	0.399	92.5	2.107	59.72	0.676
05.feb	86.66	1.036	35.55	1.144	509.54	1.547	1143.20	0.56	30.65	0.845	37.31	1.432	35.95	1.27	120.72	0.325	91.96	2.141	61	0.676
06.feb	88.12	0.999	35.91	1.144	512.51	1.598	1159.96	0.48	30.84	0.926	37.73	1.495	36.72	1.382	121.98	0.463	93.83	2.18	61.34	0.547
07.feb	89.4	0.971	36.29	1.144	519.68	1.563	1177.44	0.48	31.22	0.844	37.65	1.425	37.19	1.409	126.31	0.382	94.87	2.07	62.88	0.422
10.feb	90.42	0.934	36.53	1.126	528.99	1.546	1172.93	0.462	31.47	0.834	38.27	1.43	37.3	1.302	126.45	0.382	94.5	2.095	63.78	0.52
11.feb	92.31	1.098	36.89	0.984	535.96	1.602	1190.18	0.538	31.88	0.902	38.34	1.437	37.84	1.32	129.43	0.497	94.96	2.205	63.97	0.62
12.feb	91.76	0.979	37.19	1.204	535.92	1.618	1186.69	0.603	31.61	0.901	38.21	1.469	38.07	1.35	128.13	0.623	96.17	2.216	65.14	0.559
13.feb	91.9	1.088	37.33	1.111	544.43	1.572	1199.90	0.571	31.7	0.821	38.35	1.375	38.42	1.373	129.5	0.363	96.11	2.118	67.04	0.559
14.feb	92.1	1.065	37.34	1.17	543.99	1.525	1202.80	0.508	31.94	0.57	38.62	1.43	37.98	1.328	130.16	0.412	96.55	2.173	68.34	0.668
18.feb	91.51	1.28	37.42	1.155	545.99	1.507	1210.88	0.491	31.88	0.795	37.18	1.385	37.97	1.287	130.63	0.412	96.56	2.094	68.66	0.765
19.feb	90.99	1.036	37.51	1.084	537.37	1.548	1202.34	0.543	31.48	0.823	36.81	1.414	37.87	1.293	128.39	0.364	96.21	2.125	67.99	0.536
20.feb	91.72	1.305	37.75	1.056	531.15	1.45	1204.11	0.555	31.55	0.935	37.01	1.446	38.27	1.333	129.56	0.373	96.92	2.15	68.48	0.595
21.feb	91.52	1.106	37.98	1.03	525.25	1.691	1203.79	0.542	31.46	0.577	36.89	1.49	38.1	1.601	128.28	0.385	97.5	2.206	68.22	0.681
24.feb	91.11	1.093	37.69	1.163	527.55	1.606	1212.51	0.506	31.99	0.913	37.21	1.466	38.14	1.288	129.59	0.392	97.32	2.188	68.77	0.539
25.feb	91.11	1.064	37.54	1.125	522.06	1.508	1220.00	0.495	31.89	0.913	37.47	1.398	38.25	1.266	126.78	0.345	96.41	2.15	68.01	0.491
26.feb	91.11	0.994	37.47	1.079	517.35	1.485	1220.17	0.488	31.99	0.873	37.57	1.351	38.5	1.292	126.61	0.319	97.2	2.254	67.82	0.407
27.feb	91.36	1.032	37.86	1.065	527.67	1.487	1219.21	0.473	32.23	0.895	37.77	1.355	38.95	1.256	128.56	0.338	96.7	2.379	69.92	0.632
28.feb	92.12	1.043	38.31	1.118	526.24	1.507	1215.65	0.484	32.11	0.863	37.9	1.369	39.11	1.284	128.92	0.344	96.97	2.32	68.63	0.489
03.mar	91.56	0.921	37.78	1.077	527.76	1.444	1202.69	0.502	31.98	0.876	37.82	1.369	38.51	1.212	128.22	0.308	96.31	2.055	67.86	0.643
04.mar	93.34	1.107	38.41	1.077	531.24	1.548	1214.91	0.747	32.69	0.889	38.01	1.384	39.41	1.282	130.23	0.41	97.02	2.126	68.6	0.543
05.mar	92.59	1.082	38.11	1.05	532.36	1.544	1218.26	0.511	32.75	0.889	38.05	1.326	39.5	1.292	128.79	0.38	96.37	2.164	68.9	0.499
06.mar	92.89	1.05	38.15	1.128	530.75	1.611	1219.61	0.538	32.46	0.814	38.17	1.326	39.46	1.328	128.86	0.392	97.6	2.199	68.92	0.521
07.mar	93.32	1.061	37.9	1.145	530.44	1.668	1214.79	0.492	32.43	0.879	38.25	1.474	38.83	1.348	128.54	0.355	97.05	2.227	68.52	0.6
10.mar	93.45	1.046	37.82	0.942	530.92	1.623	1211.57	0.578	32.39	0.831	38.35	1.456	38.86	1.395	126.89	0.359	96.75	2.245	68.04	0.428
11. mar.	93.49	1.133	38.02	1.116	536.09	1.635	1199.99	0.52	32.42	0.873	38.49	1.508	38.9	1.357	125.67	0.354	96.84	1.977	67.47	0.512

HPQ	HPQ.GG	FDX	FDX.HP	CBS	CBS3825072	SXL	ETP3669496	AYR	AYR3854740	BBY	BBY4029053	HNT	HNT.GB	ADT	ADT3991009	MGM	MGM3669531	CAR	CAR3983714
27.53	2.243	139.62	2.604	63.14	3.61	72.58	2.404	18.61	2.92	40.24	3.668	29.36	3.535	39.42	2.557	23.72	2.68	40.45	5.835
28.2	1.888	139.9	2.747	63.03	3.302	72.57	2.781	18.56	2.92	40.41	3.448	29.60	3.535	39.3	2.718	23.45	2.74	40.53	5.82
28.15	1.781	138.57	2.596	62.9	3.295	72.84	2.479	18.41	2.798	39.15	3.795	29.40	3.686	39.45	2.624	23.48	2.64	40.05	5.853
28,05	1.845	140.49	2.599	63.27	2.977	73.05	1.69	18.35	2.682	38.13	3.765	30.42	3.46	39.99	2.625	24.51	2.82	39.81	5.803
27.32	1.817	140.22	2.777	62.63	3.136	72.77	2.528	18.31	2.673	37.59	3.769	30.81	3.528	39.26	2.483	24.73	2.50	40.10	5.516
27.48	1.812	141.36	2.444	62.38	3.448	72.92	2.194	18.27	2.67	37.28	3.88	31.06	3.52	39.02	2.55	24.98	2.81	40.30	5.79
27,57	1.831	142.47	2.458	62.88	3.812	72.93	2.516	18.62	2.592	37.81	3.808	30.51	3.392	38.88	2.55	25.36	2.80	40.54	5.765
27.99	1.831	140.34	2.482	60.83	3.339	73.62	2.104	18.54	2.569	36.86	3.499	30.17	3.152	38.81	2.532	25.25	2.64	39.74	5.569
28.71	1.619	142.54	2.505	60.54	3.341	73.35	2.11	18.81	2.661	36.8	3.79	31.52	3.37	38.74	2.596	25.78	2.60	40.68	5.743
28.7	1.623	142.39	2.505	60.89	3.131	73.71	2.397	18.81	2.716	37.32	3.724	32.78	3.526	38.75	2.552	25.72	2.71	41.71	5.498
29.42	1.592	141.66	2.488	60.69	3.131	73.76	2.005	18.75	2.957	26.65	4.324	32.50	3.226	38.99	2.552	25.78	2.56	41.52	5.723
29.66	1.525	140.36	2.456	60.38	3.376	73.94	1.846	18.67	2.659	24.27	4.395	32.89	3.507	39.46	2.552	26.41	2.62	40.96	5.553
29.76	1.526	141.99	2.488	59.59	3.263	74.39	1.862	19.28	2.677	24.31	4.326	33.24	3.574	39.38	2.571	26.36	2.88	40.75	5.531
29.7	1.733	142.14	2.488	60.66	3.435	74.29	1.723	19.27	2.591	25.84	4.238	33.28	3.498	39.36	2.53	25.74	2.87	41.15	5.747
29.23	1.603	140.09	2.535	59.59	3.435	74.86	1.817	18.98	2.634	25.9	4.322	34.73	3.458	38.82	2.496	25.18	2.58	40.07	5.738
28.35	1.588	134.43	2.535	58.29	3.31	74.74	2.406	18.52	2.936	24.86	4.413	33.60	3.456	38.58	2.63	23.78	2.68	38.60	5.855
28.46	1.639	131.81	2.411	58.08	3.084	73.61	2.083	18.54	2.814	24.44	4.261	33.22	3.358	38.14	2.554	23.78	2.63	38.11	5.928
28.86	1.584	134.19	2.427	58.6	3.281	74.88	2.261	18.8	2.869	24.56	5.189	32.87	3.358	38.28	2.456	24.05	2.65	38.35	5.819
28.88	1.529	131.71	2.503	57.62	3.195	75.28	1.878	18.63	3.071	23.81	4.29	32.52	3.417	37.81	2.534	23.28	2.76	37.49	5.891
29,11	1.547	133.62	2.385	59.06	3.195	75.94	1.649	18.8	2.798	22.57	4.365	33.07	3.513	31.4	2.708	24.11	2.65	38.27	5.855
28.86	1.69	133.17	2.138	58.62	2.832	77.91	1.891	18.7	2.757	23.39	4.18	32.89	3.268	30.04	2.67	24.36	2.67	37.71	5.922
27.91	1.424	129.3	2.362	56.21	3.102	78.18	1.833	18,04	2.832	22.92	4.258	31.93	3.479	28.83	2.571	23.86	2.65	35.92	5.939
28.19	1.46	130.28	2.304	57.77	3.174	76.01	2.107	18.33	2.789	22.63	3.916	32.25	3.476	30.2	2.679	24.20	3.06	37.81	6.001
27.88	1.494	129.83	2.304	58.84	3.167	76.86	2.934	18,07	2.789	22.97	4.026	31.86	3.585	29.75	2.632	23.67	3.07	37.63	5.965
28.35	1.36	130.7	2.41	60.18	3.215	78	2.018	18.43	2.854	23.56	3.894	32.09	3.315	29.96	2.632	24.29	2.79	38.82	5.929
28.93	1.396	131.62	2.212	60.39	3.162	79.58	2.185	18.62	2.851	24.39	4.225	31.73	3.313	31.53	2.763	24.80	2.79	38.95	5.655
28.67	1.43	129.74	2.585	60.48	3.099	79.85	2.365	18.34	2.848	25,08	4.255	32.75	3.494	30.93	2.712	24.64	2.95	39.25	5.712
29.22	1.341	133	2.426	60.79	3.15	80.72	2.623	17.64	2.845	25.25	4.224	32.52	3.476	30.71	2.634	25.25	2.70	39.31	5.777
29.28	1.425	132.85	2.595	61.74	3.288	81.16	2.405	17.67	2.911	24.73	3.895	31.39	3.468	30.55	2.25	25.32	2.95	38.68	5.676
29.69	1.365	132.92	2.406	64.5	3.072	82.04	1.848	18,04	2.733	24.52	4.228	31.74	3.468	31.3	2.493	25.58	2.63	38.67	5.677
29.88	1.275	133.77	2.454	64.84	3.205	81.8	2.378	17.92	2.749	24.62	4	32.41	3.448	31.61	2.327	26.02	2.50	38.75	5.584
29.44	1.285	132.64	2.377	66.17	3.156	82.7	2.277	18.28	2.902	24.83	4.038	32.13	3.368	31.3	2.497	25.85	2.66	38.92	5.482
29.31	1.339	131.21	1.953	65.7	3.047	82.09	2.434	18.21	2.893	24.72	4.067	31.66	3.368	31.45	2.412	25.75	2.67	40.00	5.57
30.05	1.335	132.84	1.953	65.79	3.216	80.96	1.894	18.51	2.874	24.68	4.097	31.63	3.499	30.86	2.454	26.83	2.98	43.61	5.534
29.65	1.499	134.75	2.234	65.94	3.245	81.25	2.035	18.51	2.769	24.48	4.188	31.82	3.389	30.57	2.361	27.05	2.56	44.88	5.243
29.78	1.499	134.72	2.342	66.34	3.236	80.25	1.954	18.89	2.884	25,01	3.6	33.33	3.529	31.27	2.482	27.84	2.44	46.05	5.159
29.74	1.245	132.69	2.274	65.64	3.197	80.55	1.897	19.82	2.645	25.14	4.187	33.00	3.368	31.7	1.959	27.54	2.55	46.64	5.098
29.76	1.198	132.23	2.317	65.42	3.168	82.29	1.89	19.87	2.769	25.65	4.154	33.55	3.368	30.9	2.444	27.49	2.60	46.75	5.033
29.61	1.137	133.4	2.317	66.28	3.129	81.69	2.042	19.67	2.846	25.4	4.092	33.48	3.455	31.34	2.374	27.86	2.69	47.89	5.37
29.74	1.393	133.18	2.292	66.96	3.129	82.74	3.052	19.5	2.662	26.46	3.973	34.05	3.33	30.71	2.444	27.55	2.81	46.99	5.246
29.59	1.163	133.23	2.436	65.83	3.129	82.08	2.175	19.35	2.935	25.98	4.06	33.64	3.451	30.48	2.411	27.60	2.52	46.84	5.285
29.98	1.17	136.94	2.265	67.3	3.125	83.52	2.042	19.63	2.734	25.63	4.06	34.08	3.464	30.69	2.366	28.29	2.39	48.39	5.2
29.8	1.127	136.76	2.316	67.26	3.185	82.45	2.456	19.7	3.003	25.37	4.125	34.35	3.464	31.24	2.563	28.22	2.41	48.77	5.2
29.95	1.246	137.14	2.227	67.43	3.106	82.59	2.233	19.94	2.761	25.27	3.753	34.66	3.252	30.93	2.485	28.39	2.41	48.66	5.242
30.2	1.246	137.43	2.948	67.4	3.415	82.95	2.149	19.9	2.837	25.63	4.149	34.71	3.32	30.85	2.456	28.29	2.45	48.34	5.419
29.9	1.289	137.86	2.939	66.76	3.415	83.07	1.821	19.96	2.872	26,07	4.18	35.09	3.317	29.98	2.563	27.71	2.40	48.31	5.499
29.91	1.385	137.29	2.317	66.09	3.182	83.94	2.184	19.58	2.791	25.8	3.708	35.10	3.317	29.73	2.532	27.05	2.48	47.07	5.327

AMD	AMD.GL	ANR	ANR4004999	PKG	PTV.GH	NQM	JNC3667148	BZH	BZH.GU	GGS	GGS.AC	VRS	VRS.GC	Risk free rate
3.95	6.054	7.26	2.226	62.9	5.94	13.32	4.871	24.71	6.773	1.61	21.056	0.65	41.263	0.07
4	6.054	7.04	2.226	62.62	5.94	13.38	4.193	24.32	6.143	1.7	19.77	0.65	41.993	0.07
4.13	5.533	6.94	3.119	61.96	5.533	13.53	4.714	23.76	6.11	1.63	18.838	3.21	31.805	0.05
4.18	5.635	6.5	4.066	63.21	5.937	13.6	4.404	22.84	6.861	1.59	18.757	4.38	24.742	0.04
4.18	5.45	6.52	3.629	62.95	5.937	13.62	4.404	22.96	6.636	1.53	18.957	3.75	24.605	0.05
4.09	4.793	6.27	4.202	63.06	5.598	13.64	5.387	22.86	6.093	1.47	17.409	3.85	23.55	0.04
4.17	4.784	6.21	3.474	64.24	5.521	13.76	4.551	23.12	6.629	1.42	18.761	4.15	23.463	0.05
4.13	5.469	6.14	4.109	64.21	5.5	13.71	5.07	22.27	3.667	1.43	18.144	3.72	25.807	0.03
4.3	5.065	6.13	4.171	65.05	5.059	13.71	4.548	22.5	6.623	1.41	17.308	3.69	26.059	0.04
4.47	4.349	6.3	3.439	65.13	5.744	13.81	4.982	22.62	6.623	1.44	19.3679	3.13	27.687	0.04
4.38	5.181	6.4	3.439	65	5.664	13.82	4.708	22.52	4.553	1.45	18.087	3.13	30.227	0.04
4.18	4.666	6.32	3.561	64.57	5.35	13.88	4.708	22.48	6.603	1.58	18.529	2.75	30.353	0.05
4.17	4.145	6.17	3.561	65.26	5.333	13.85	4.708	22.94	5.119	1.59	17.406	2.7	29.825	0.04
3.67	5.116	6.25	3.775	65.84	5.333	13.93	4.232	23.67	6.731	1.64	17.237	2.72	27.07	0.04
3.62	5.662	6.2	3.909	64.93	5.263	13.94	4.755	23.49	4.92	1.53	17.603	2.68	27.284	0.04
3.47	5.608	5.83	4.829	62.16	5.463	13.98	4.755	21.7	6.726	1.5	19.836	2.89	26.795	0.04
3.41	6.132	5.71	4.821	61.84	5.325	14.09	4.755	21.47	6.27	1.42	21.554	2.53	28.473	0.05
3.54	5.223	5.84	4.821	62.66	5.315	14.1	3.917	22.57	6.27	1.47	20.772	2.76	28.493	0.05
3.48	5.806	5.82	4.821	63.49	5.315	14.03	5.022	22.76	5.008	1.44	20.398	2.65	26.844	0.04
3.48	5.799	5.85	4.81	63.63	5.303	14.01	4.845	21.96	6.71	1.47	20.277	2.99	25.261	0.02
3.43	6.114	5.68	5.009	64.24	5.248	14.07	5.332	22.51	6.707	1.48	18.758	3	25.347	0.02
3.33	5.313	5.22	5.009	62.02	5.246	14.09	4.856	21.53	5.742	1.47	20.291	2.88	26.11	0.05
3.37	5.142	5.36	5.993	62.68	5.668	14.04	4.842	21.34	6.903	1.38	20.557	3.06	25.318	0.06
3.31	5.278	5.09	6.308	62.53	5.505	13.99	4.838	20.77	6.691	1.37	20.579	2.98	25.751	0.07
3.41	5.65	5.15	6.401	62.83	6.435	13.93	4.838	21.25	4.904	1.33	20.182	3.08	26.609	0.07
3.47	5.156	5.3	6.168	64.65	5.482	13.98	4.836	21.2	5.512	1.38	20.279	2.79	26.501	0.08
3.63	5.265	5.09	6.661	65.15	5.004	14	4.967	21.37	5.501	1.39	19.072	2.8	25.911	0.07
3.7	5.256	5.26	6.42	65.19	5.151	14	4.834	21.24	5.489	1.43	20.252	2.82	25.092	0.05
3.69	5.256	5.06	6.579	71.66	5.502	13.88	4.02	20.66	5.132	1.39	19.611	2.73	25.158	0.05
3.7	3.598	5.17	6.542	70.94	5.991	13.97	3.52	20.95	4.665	1.57	20.646	2.53	24.196	0.03
3.69	4.75	5.06	6.756	72.19	5.452	13.97	4.828	21.26	4.455	1.49	20.892	2.58	26.058	0.02
3.7	5.072	5.11	6.248	72.16	5.214	13.94	4.91	20.78	5.407	1.54	20.246	2.51	24.608	0.05
3.72	5.104	5.25	6.274	71.66	5.207	14.01	4.657	20.57	5.219	1.44	19.149	2.41	24.608	0.06
3.69	5.157	5.18	6.237	71.93	5.205	14.02	3.827	20.78	5.219	1.45	17.57	2.45	24.43	0.05
3.69	5.157	5.19	6.277	71.82	5.346	14.06	4.599	21.49	6.641	1.47	18.247	2.42	28.819	0.05
3.71	5.662	5.27	6.298	71.89	5.122	14.03	4.599	21.95	3.353	1.5	17.135	2.36	26.008	0.05
3.69	4.355	5.19	6.239	71.53	5.12	14.04	4.617	22.3	3.498	1.44	19.528	2.36	26.008	0.05
3.7	4.607	5.34	6.289	71.53	4.793	14.02	5.155	23.49	3.906	1.39	18.34	2.34	26.008	0.05
3.71	4.753	5.58	5.41	71.86	4.802	14.01	3.819	23.5	5.259	1.41	18.775	2.35	25.707	0.04
3.71	4.742	5.37	6.075	72.49	5.064	14.09	3.582	23.19	4.04	1.4	18.66	2.45	25.232	0.05
3.67	5.45	5.32	5.981	72.46	4.918	14.09	4.552	23.1	4.194	1.37	18.853	2.74	27.987	0.05
3.7	4.401	5.31	6.211	73.87	4.705	14.12	3.735	23.4	3.832	1.48	18.12	2.7	28.149	0.05
3.71	5.987	5.24	6.211	74.2	4.307	14.14	4.618	23.21	3.601	1.44	20.285	2.83	28.643	0.06
3.73	6.889	5.48	5.99	73.79	4.516	14.02	4.431	22.99	3.756	1.43	18.022	2.41	27.874	0.05
3.95	4.177	4.8	7.531	73.8	4.894	13.89	4.431	22.35	1.863	1.41	17.679	2.28	26.4	0.06
3.81	4.651	4.63	8.192	73.44	5.022	13.91	4.431	21.36	4.779	1.41	19.277	2.1	26.697	0.05
3.85	4.692	4.41	7.768	72.38	4.783	13.97	3.781	21.25	4.406	1.37	18.531	2.02	28.377	0.05

Appendix 2. Scripts for Shazam (software)

```
*###############################
*Basic regressions
*###############################
*1ø. Cochrane- Orcutt regression on yield
and stock price
sample 1 47
?auto Y S / coef=Beta TRATIO=TR rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
*###############################
*2ø. Cochrane- Orcutt regression on
change in yield and stock return
sample 1 47
genr rS = (S-lag(S))/lag(S)
genr rY = (Y-lag(Y))/lag(Y)
genr dy = y-lag(y)
genr ds = s-lag(s)
sample 2 47
?auto dY rS / coef=Beta TRATIO=TR rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
*###############################
*Regressions to investigate various time
lags.
*###############################
*3ø. Cochrane- Orutt regression on change
in yield.
*Leading and contemperaneous stock
return as independent
*variables.
sample 3 46
?auto dY rS(-1.0) / coef=Beta
TRATIO=TR rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
```

```
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
*###############################
*4ø. Cochrane- Orutt regression on change
in yield.
*Contemperaneous and lagged stock return
as independent
*variables.
sample 3 46
?auto dY rS(0.1)/ coef=Beta TRATIO=TR
rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
*###############################
*5ø. Cochrane- Orutt regression on change
in yield. Leading,

contemperaneous
* and lagged stock return as independent
variables.
Sample 3 46
auto dY rS(-1.1) / rstat
*###############################
*"Do-loops"
*###############################
*6ø. Loop to find the time relationship
between dY and rS
*Generating the lagges of dY
do #=0,3
sample 5 47
genr LdY# = Lag(dY,#)
endo
*
do #=0,3
sample 5 47
?auto LdY# rS / coef=Beta TRATIO=TR
rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
```

```
gen1 laic#=$laic
gen1 aic#=$aic
endo
print laic0-laic3
print aic0-aic3
*###############################
*7ø. The same procedure, but now I have
lagged dS
do #=0,3
sample 5 47
genr LrS# = Lag(rS,#)
endo
*
do #=0,3
sample 5 47
?auto dY LrS#/ coef=Beta TRATIO=TR
rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
gen1 laic#=$laic
gen1 aic#=$aic
endo
print laic0-laic3
print aic0-aic3
*###############################
*8ø. Checking 1-lag rS to veriy the
letterature who
*suggests that this is true
sample 3 47
?auto dY lrS1 / coef=Beta TRATIO=TR
rstat
GEN1 DF1=$DF
SAMPLE 1 $k
GENR TRA=ABS(TR)
DISTRIB TRA / TYPE=T DF=DF1
CDF=CDF1
GENR PVAL2=2*(1-CDF1)
print $R2 $dw Beta pval2
*###############################
*9ø. Checking multicolinearity
sample 6 47
stat lrS0-lrS3 / pcor
stat ldY0-ldY3 / pcor
*###############################
*miscellaneous correlations

sample 1 47
stat Y S / pcor
sample 2 47
stat rY
stat dY rS / pcor
stat dY S / pcor
stat Y rS / pcor
sample 3 47
stat LrS1 dY / pcor
stat ldY1 rS / pcor
stat ds dy
*###############################
*Tests for unit root and cointegration
sample 1 47
coint y s
sample 2 47
coint y s / NDIFF=1
sample 1 47
coint y s / type=resd
graph y / lineonly
graph s / lineonly
genr rS = (S-lag(S))/lag(S)
genr dy = y-lag(y)
genr ldy = lag(dy)
genr ly = lag(y)
sample 3 47
OLS dy ldy rs / rstat dlag
sample 2 47
OLS y ly s / rstat dlag
graph dy
```