BI Norwegian Business School - campus Oslo

# GRA 19703

## Master Thesis

Thesis Master of Science

Time-varying stock market return predictability: Do we have what it takes?

Navn: Ousman Touray, Ole Arne Jakobsen Steen

Start: 15.01.2020 09.00

Finish: 01.09.2020 12.00

# TIME-VARYING STOCK MARKET RETURN PREDICTABILITY: DO WE HAVE WHAT IT TAKES?

Master Thesis

by

Ole Arne Jakobsen Steen & Ousman Touray

*MSc in Business with Major Finance*

SUPERVISOR:

**Patrick Konermann**

Oslo, August 31, 2020

## ABSTRACT

We use a dividend-yield model from Campbell and Shiller (1988) to forecast the future stock market return on the U.S and Norwegian data from 1984-2018. We use the method from Cochrane (2008), by regressing a Vector Autoregression (VAR)-system and check for forecasting power in the long-run. We find that return gives *stronger* evidence against unforecastable null-hypothesis for return in the U.S data than the Norwegian data. Norwegian market gives *stronger* evidence for the dividend growth. $R^2$ increases in the long-run for dividend growth in the Norwegian data, while $R^2$ decreases for return. The opposite appears for the U.S data. We conclude that stock market predictability using the dividend yield model from Campbell and Shiller (1988) and Cochrane (2008) method gives *different* results for Norwegian data compared to the U.S data.

# Acknowledgements

This Master thesis marks the end of our study programme MSc in Business with Major in Finance at BI Norwegian Business School. We thank our Supervisor Patrick Konermann for excellent guidance on how to proceed in our Thesis. We also thank John H. Cochrane for additional comments on his data and methodology. We also thank Goyal for informative replies to our concerns. We thank BI Norwegian Business School for excellent two years of study. At last, we thank our families and friends for supporting during these two years and during our Master thesis.

# Table of Content

# List of Abbreviations

**AR** Autoregression

**VAR** Vector Autoregression

## List of Figures

# List of Tables

# List of Symbols

| | |
|---|---|
| $\beta_d$ | One-period beta coefficient for dividend-growth |
| $b_d$ | One-period beta coefficient for dividend-growth |
| $b_d^{lr}$ | Long-run beta coefficient for dividend-growth |
| $\beta_r$ | One-period beta coefficient for return |
| $b_r$ | One-period beta coefficient for return |
| $b_r^{lr}$ | Long-run beta coefficient for return |
| $\Delta d_{t+1}$ | The dividend-growth (log of the change in dividend pay-out from time t to time t+1) |
| $dp_t$ | The dividend-yield at time t |
| $d_{t+1}$ | Dividend at time t+1 |
| $\phi$ | The autocorrelation function for dividend-yield |
| $p_{t+1}$ | Price at time t+1 |
| $\rho$ | A point estimate of $log[1 + e^{(p_{t+1} - d_{t+1})}]$ |
| $r_{t+1}$ | Return at time t+1 |
| $t_d$ | Test statistic for dividend-growth |
| $t_r$ | Test-statistic for return |

# 1 Introduction

Our thesis aims to examine Cochrane's VAR system (2008) as applied to the dividend-yield model of Campbell and Shiller (1988) and to assess whether or not the system is applicable to predict return in the U.S and the Norwegian stock market. How can this model be used to forecast the stock market return, and how much of the variation in return and dividend growth can be explained by dividend-yield (DP-ratio) for both of these markets? This particular system (which we will henceforth refer to as the Cochrane System) uses the dp-ratio of today as the dependent variable to forecast the one-year and long-run return, dividend-growth and dividend-yield. Generally, a dividend is a distribution of the company's cashflows that is paid out to the shareholder (also known as a stockholder) (Chen (2019)). Because of this, higher dividends signal higher earnings for the company, which naturally involves a higher expected return for the shareholders. Due to the structure of market mechanics, the price of a stock is adjusted according to the dividend payout. Hence, a high dividend payout gives a low ex-post stock price. This knowledge about the way dividends affect the stock price is what we will come to refer to throughout our paper as the *economic intuition*. Any true knowledge that can explain what actually moves prices is vital for any participant or spectator of the stock market. Researchers, institutional investors, analysts, among others, would, with such knowledge, be able to construct holding- or trading strategies to generate profit from the market and improve the asset allocation process in the pursuit of capital gains.

The S&P 500, NASDAQ and Dow Jones are some of the indexes that are being used as a standard benchmark the performance of the stock market in the

U.S. (Tradingview (2020)). For the Norwegian Stock market, the performance is measured by the level of the Oslo Stock Exchange (2020). Today, we are facing a great deal of unclarity with regarding the most dominant factor(s) that can predict return. How can this be when we have numerous researchers have tried to identify parameters that can explain all variations in the stock market with different models? Some examples are CAPM (Sharpe (1964)) and the respective three- and five-factor model (Fama and French (2015)), and there are many more. As of today, there is no common approach that can perfectly capture and explain all market variation in-sample, and especially out-of-sample. Moreover, the Cochrane System (2008), which has been used with reasonable accuracy to forecast the U.S. Market, has given too inconsistent results when applied to European markets. For instance, this can be seen in Engsted and Pedersen (2010) and Monteiro (2018), who arrived at completely different conclusions for the European market compared to the U.S. Market, and even found inconsistencies between different European markets, despite always using the same predicting variable. We will be examining Cochrane's System further, especially with respect to adequacy for the Norwegian market.

To summarize our empirical procedure, we used the dividend-growth model of Campbell and Shiller (1988) and the Cochrane system (2008) for the U.S. and Norwegian stock market from 1984 to 2018. We generated results from a one-period regression, in addition to long-run regression to forecast for a maximum of 11 years, given our current data. Based on the *economic intuition*, we expect a negative coefficient for the DP-ratio on dividend-growth. From the short-term (1 year) regression, this only seems to apply to the Norwegian market, and not the U.S. market. The long-run forecasts imply an increasing return-coefficient as we increase the number of lags in the U.S. This is what we would expect based on the fact that the stock market has (historically speaking) yielded positive returns in the long-run. However, when it comes to the Norwegian stock market, the long-run forecast implies a lower return by generating negative return coefficients - which is not expected to see. The estimates for return and dividend-growth are calculated using a joint hypothesis test, assuming that one dependent variable is forecastable, while the other is not. The coefficients for return are based on the null-hypothesis that return is not forecastable, and the alternative-hypothesis that dividend-growth is forecastable. By specifying each test this way, we obtain more consistent evidence against a null-hypothesis that assumes that dividend-growth is not forecastable, rather than a null-hypothesis that assumes that return is. This is the case for both the U.S. and the European markets. In order to conduct a more accurate out-of-sample (OOS) test, we believe that it would be necessary with more data than we have. Hence, we have only conducted in-sample (IS) tests for long-run forecasting.

We start our paper by presenting some of the most relevant literature. We give a brief description of each paper's findings, why they are essential, and how the different papers are connected. We then present our empirical results and talk about distribution, volatility, and the correlation between the different variables. The rest of the paper will include an analysis of the short- and long-run regression of the real- and excess return and dividend-growth on the DP-ratio and the time-varying probabilities for the dependant variable. Finally, we will summarize and give an overall presentation of how our findings are interconnected and related to the literature.

## 2 Literature Review

In this thesis, we are interested in looking at stock return predictability. We know that investors yield capital gains from stock either through cash flows (e.g., dividends) or an increase in stock price, or both (Investopedia (2020)). This means the return is dependent on these two elements (among others). Therefore, we can form an equation where the investor buys a stock at time t, earns cash flow (dividends) at time t+1, and sells the stock at time t+1. Starting off, we can write the return of a stock, portfolio, or index as:

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}$$

$$R_{t+1} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}})\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}}$$

$$r_{t+1} = log[1 + e^{(p_{t+1} - d_{t+1})}] + \Delta d_{t+1} + (d_t - p_t)$$

Where $R_{t+1}$ is the return at time t+1 and the equivalent written in lowercase letters is the log of the corresponding variable.

### 2.1 Campbell and Shiller Decomposition

This equation is the foundation of the dividend-yield model which was presented by Campbell and Shiller (1988). According to them, the dividend-to-price ratio is often interpreted as reflecting the prospect of future dividends. For instance, if the DP-ratio is high, forecasts should imply a lower expected future dividend growth due to a low price (Campbell and Shiller (1988)). Conversely, Campbell and Shiller's alternative interpretation is that the DP-ratio reflects the amount of future dividends that are discounted from the price as it is today. In that sense, it is like a discount factor. At that time, the literature lacked a more comprehensive and thorough analysis of time-varying DP-ratio with respect to the dividend-growth and the discount factor.

In their article, Campbell and Shiller approximate the part within the square bracket in the equation above, $log[1 + e^{(p_{t+1} - d_{t+1})}]$ using the Taylor Approximation, which is shown in the appendix. This term is seen as the way the change in price-dividend ratio affects the return at time t+1, using a point estimation. The point estimation can, therefore, be used to approximate the long-run estimates. Hence we see that, when using this same approximation, both single period returns and long-run returns should be affected by the DP-ratio, dividend growth, and the point estimate of the the price-dividend ratio.

### 2.2 Fama and French forecasting power

Alongside Campbell and Shiller's article, the year 1988 also saw the publication of the better known Fama and French article (1988). In this article Fama and French look at predicting stock return by using the dp-ratio as the independent variable. That is, they look at the regression $r(t, t+T) = \alpha + \beta(\frac{D_t}{P_t}) + \varepsilon(t, t+T)$, where $r$ is the return and $\frac{D_t}{P_t}$ is the dividend yield at time t. Moreover, they find that the $R^2$ for return increases as they increase the number of years that the model is forecasting for. An interesting takeaway from their results is that dividend yield tends to explain more of the expected return variances when regressing more than five-year returns, compared to one-year returns. Thus, long-run regressions lead to higher explanatory power on dividend yield and account better for the variation on return. To quote from the Fama and French article directly: "[...] high autocorrelation causes the variance of expected returns to grow faster than the return horizon" Fama and French (1988, p. 1). Hence, their findings are relevant to our findings and helps us to interpret results that can validate or challenge the work of Campbell and Shiller (1988).

### 2.3 Forecasting articles in 2005-2018

During the period 2005-2008, many articles appeared on the topic of forecasting, including those written by Cochrane (2008), Welch and Goyal (2008), and Campbell and Thompson (2008). In the article from Cochrane (2008), he approached the topic of return forecasting in the U.S. with a first-order Vector Autoregression (VAR)-system of log return, log dividend growth and log of future dividend-yield, by using the dp-ratio as the predicting variable. The equation is written as follows:

$$r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1} \tag{1}$$

$$\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1} \tag{2}$$

$$d_{t+1} - p_{t+1} = \alpha_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1} \tag{3}$$

The relationship ($b^{lr}_r - b^{lr}_d = 1$), which is obtained from the VAR system above, is mainly how Cochrane specifies the null-hypothesis for the unpredictable-return and predictable dividend growth, respectively. Setting a null hypothesis in this way assumes that return is unforecastable, while dividend-growth is not forecastable. Alternatively, we can set up a null hypothesis that assumes that dividend growth is not forecastable while return is forecastable. However, it is important to note that return and dividend-growth cannot both be set as forecastable, nor is it possible to set neither as unforecastable using this system. The dividend yield model developed by Campbell and Shiller (1988), sets the foundation for the empirical procedure used here in Cochrane's article. The long-run and one-period regressions are used for the S&P 500 to examine the explanatory variable in one-period vs long-run forecasts, as well as assessing the way the estimates move compared to the historical mean, which is used as a benchmark for our expectation. Cochrane (2008) finds that a null-hypothesis that assumes that dividend growth is forecastable gives stronger evidence against the null-hypothesis that assumes that return is unforecastable.

In their 2008 article disagreeing with Cochrane (2008), Welch and Goyal regress the equity premium on the market using different variables. Among the different variables such as interest rates, earnings price ratio, consumption and wealth, they also used the DP-ratio. They take a more offensive position in forecast ability and disagree with Cochrane's approach, which finds that dividend yield must predict returns if it fails to predict dividend growth. They conclude that "He has strong priors, placing full faith in a stationary specification of the underlying model—even though Welch and Goyal (2008) have documented dramatic increases in the autocorrelation of dividend growth" (Welch and Goyal (2008, p. 1505)). They also conclude that most models fail to beat the unconditional benchmark (the historical market return) and underperform when it comes to their ability to forecast. Similarly to Cochrane's approach, Welch and Goyal generated the probability distribution from the bootstrapped distribution. This is described in more detail in Data and methodology.

In 2010, the researchers Engsted and Pedersen used the Cochrane System (2008) to look at predictability for U.S. and European data. The article reveals major inconsistencies when applying the Cochrane system on the U.S.- and the European markets. In their article, the $R^2$ for dividend growth in Sweden and Denmark is higher than it is for return. The article uses the same methodology as Cochrane, using the same number of lags: up to k=20. Specifically, the $R^2$ for return in Sweden is only 3.44 for 20 years, while it increases to as much as 45.10 in the U.S. data. For dividend growth, Sweden obtains a $R^2$ of 0.366, whereas it is as low as 0.95 for the U.S. data. The probability values from the Monte Carlo Simulation confirm that, in terms of the Swedish market, there is better evidence against the unforecastable null for dividend growth than unforecastable null for the return. The opposite it true for the U.S. data; "In addition, in Sweden and Denmark dividend growth is strongly predictable by the dividend–price ratio in the 'right' direction while returns are not predictable" Engsted and Pedersen (Engsted and Pedersen (2010, p. 587)). Monteiro's article of 2018 found the same pattern. Dividend growth was shown to be better forecastable than the return for Spain and Italy, while showing the opposite for the U.S. and the U.K. market. Monteiro (2018) also finds that in France only returns are predictable, while in Germany there is evidence for both return and dividend growth predictability. Furthermore, when using Japanese data neither returns nor dividend growth are forecastable. They conclude as follows in the paper: "Generally, there is no clear connection between dividend smoothness and predictability" (Monteiro (2018, p. 1)). Key takeaways from the previous papers are the way $R^2$, bootstrap distributions, and probabilities is used to examine how the model performs in long- and one-period on return and what the result implies for the respective hypotheses.

### 2.4 Differences & similarities

So far, we have shed light on the connection between Campbell and Shiller (1988) and the articles that came later. Cochrane's article (2008) revealed exciting methods and findings, such as the VAR system and the better long-run predictability. We also found it interesting that there is stronger evidence against the forecastable null for dividend growth than the unforecastable null for return. In the Methodology we describe how we applied Cochrane's method to explore this null hypothesis system for the Norwegian data. Campbell and Shiller's approximation and dividend price model (1988) is used by Cochrane (2008) and sets the baseline for our empirical procedure.

The way we specify return regression in our paper is identical to how Fama and French specify it in their 1988 article. We also include Cochrane's VAR system as applied to Campbell and Shiller (1988). Fama and French (1988) examined how $R^2$ performs for a longer horizon. We also did this, as well as including the power of forecastability by bootstrapping and Monte Carlo Simulation. Fama and French (1988) conclude that dividend yield does not explain much on the return variation for one period, which is similar to Cochrane (2008). It is worth pointing out that Welch and Goyal (2008) use a different method whereby they regress different variables instead of choosing only one variable or a VAR system. While Campbell and Shiller (1988) and Cochrane (2008) regress with one independent variable, Welch and Goyal (2008) use 10. Besides, they perform an out-of-sample (OOS) sample test, which we do not include. The data from the literature above comes from NYSE in the U.S., while our data comes from the Oslo Stock Exchange in Norway. We are using data from 1984- 2018, while most of the data in the literature mentioned comes from 1926-2004, except for Welch and Goyal (2008) who use data that goes as far back as 1871.

These are some of the similarities and differences between our paper's methods and the methods applied in the literature we have studied. It is also worth noting that Engsted and Pedersen (2010) found that dividend growth gives better predictability than return for the European market, compared to the U.S. market. However, their results imply lower long-run estimates for return for all European countries, except for the U.K. This is different from the article from Cochrane (2008) who found better predictability for the return in the long-run. The same applies to Fama and French (1988), who found that $R^2$ increases for the return in the long-run. On the whole, the empirical approaches of Campbell and Shiller (1988), Fama and French (1988), Cochrane (2008), Engsted and

Pedersen (2010), and Monteiro (2018) are all interconnected and show more similarities than differences (Welch and Goyal (2008) are the exception who appear to have had the most different approach out of all of these papers).

## 3 Methodology and hypotheses

In this paper, we followed the methodology outlined in Cochrane's article (2008). Below is a numerated list of expressions, which we will refer to as *identities*. The derivation of the steps and calculations revolving the underlying Vector Autoregression (VAR)-system, linearization of return, long-run regressions, and the Monte Carlo simulations are described in depth in appendix. As briefly mentioned, the article uses the first-order Vector Autoregression (VAR) system, which is defined as follows:

$$r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1} \tag{1}$$

$$\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1} \tag{2}$$

$$d_{t+1} - p_{t+1} = \alpha_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1} \tag{3}$$

where $r_{t+1}$ is the log return, $\Delta_{t+1}$ is the log dividend change, and $d_{t+1} - p_{t+1}$ is the difference between log dividend and log price at time $_{t+1}$, hereafter dividend yield at $_{t+1}$. Campbell and Shiller (1988) linearize, $r_{t+1}$;

$$r_{t+1} = \rho(p_{t+1} - d_{t+1}) + \Delta_{t+1} - (p_t - d_t) \tag{4}$$

From this linearization, the identities for coefficients and errors are obtained in Cochrane (2008) and is defined as follows:

$$\beta_r = 1 - \rho\phi + \beta_d \tag{5}$$

$$\varepsilon^r_{t+1} = \varepsilon^d_{t+1} - \rho\varepsilon^{dp}_{t+1} \tag{6}$$

Using the identity (5), we can form the hypothesis that return is unforecastable (equation 7) and dividend growth is forecastable for *one-period regression*:

$$
\begin{aligned}
H_0 &: b_r = 0, b_d = \rho\phi - 1 \\
H_A &: b_r = 1 - \rho\phi, b_d = 0
\end{aligned} \tag{7}
$$

$H_0$ assumes that return is **unforecastable** and dividend-growth is **forecastable**. Alternative hypothesis assumes the opposite.

Note that we only use one hypothesis in the one-period regression, and two for the long-run forecasts. The reason to why is that we are looking at the *variation* in the long-run for return *and* dividend-growth. Further, we obtained the OLS estimates by running the VAR system in (1)-(3). The estimates were calculated in a Monte Carlo simulation to check for probabilities that the coefficients are greater than the simulated sample value under the null hypothesis. The probabilities indicate how likely it is for the sample coefficients to appear by pure chance (Cochrane (2008)). The same is done with the t-statistics. In his article Cochrane (2008) divides identity (5) by $1 - \rho\phi$ to obtain,

$$
b_r^{lr} - b_d^{lr} = 1 \tag{8}
$$

where **lr** denotes the long-run estimate of the corresponding coefficient. Using this method, we defined our hypotheses of the return being unforecastable (equation 9) and dividend growth being unforecastable (equation 10) for *ong-run*l:

$$\begin{aligned} H_0 &: b_r^{lr} = 0, b_d^{lr} = -1 \\ H_A &: b_r^{lr} = 1, b_d^{lr} = 0 \end{aligned} \tag{9}$$

$H_0$ assumes that return is **unforecastable** and dividend-growth is **forecastable**. Alternative hypothesis assumes the opposite.

and

$$\begin{aligned} H_0 &: b_r^{lr} = 1, b_d^{lr} = 0 \\ H_A &: b_r^{lr} = 0, b_d^{lr} = -1 \end{aligned} \tag{10}$$

$H_0$ assumes that dividend-growth is **unforecastable** and return is **forecastable**. Alternative hypothesis assumes the opposite.

The full derivation of (8) is described in appendix D.3 long-run. The Monte Carlo simulation obtains the long-run estimates and simulated probability values. By applying identity (8) we are able to distinguish how much of the variation on dividend yield is caused by the return and how much is caused by the dividend growth. Additionally, we examined statistics and probabilities for $\phi$, which Cochrane defines as the autocorrelation function on dividend growth (Cochrane (2008)). These probabilities represent the effects on $b_r$ and $b_d$ when the autocorrelation on dividend-yield $\phi$ increases. The direct and indirect estimates were calculated using weighted and unweighted regression coefficients. The probabilites are likelihood of rejecting the $H_0$. Hence, Lower probability values indicates stronger evidence to reject the null-hypotheses. For instance, if the probability values for an unforecastable null-hypothesis for return is low, this means that we find stronger evidence to reject the null-hypothesis that return is not foreastable. In the long-run, this indicates indicates that most variation in dp-ratio comes from return, since we would now believe the alternative hypothesis.

Something important to note is that we have checked if the parameter estimates from an OLS regression are BLUE. We conducted the appropriate tests for all assumptions in all of the datasets, which can be seen in Appendix C-Assumptions. In short, all of the dependent variables shows signs of positive autocorrelation in the residuals, which implies a great deal of heteroscedasticity in all samples for the U.S. and the Norwegian market from an Durbin-Watson test. The estimates that were tested are derived from an OLS regression. The use of generalized method of moments (GMM) in the OLS allows us to account for the serial autocorrelation in the error-term, which would result to more heteroscedasticity in the standard errors (Hansen (1982)). In addition, since we are following the method of Cochrane (2008), we use an OLSGMM and not an OLS regression. The use of the method was validated by the script Cochrane has available on his website. Finally, We use the delta-method for standard errors.

## 4 Data description

As previously stated, we are using the Cochrane System (2008) in this thesis. We used data from Cochrane's article to replicate the outputs from his paper. The U.S. data from 1926-2004, as well as the value-weighted return on NYSE with and without dividend, which Cochrane used in his 2008 paper, is available on his website (Cochrane (2020)). The dividends are not directly observable; they are distinguished by subtracting the return with dividends from the return without dividends (vwretd-vwretx). This calculation was essential in order to set up the dividend-yield at time t as the independent variable ($d_t - p_t$), as well as dividend-growth ($\Delta d_{t+1}$), return ($r_{t+1}$) and dividend-yield at time t+1 ($dp_{t+1}$) as the dependent variables. The risk-free rate and Consumer Price Index (CPI) are the three-month treasury bill and the CPI from 1926-2004, respectively (collected from Cochrane (2020)).

When comparing the Norwegian and U.S. data from 1984-2018, we used additional sources to gather the U.S. data. Due to restrictions on more recent market data on CRSP, we could not retrieve matching, nor additional data for the years after 2004. Fortunately Welch and Goyal (2008) have annually updated files on their web page, which made it possible to obtain an extended sample of the value-weighted return for the U.S and the t-bill. To ensure that the data was the same, we checked that there was a correlation between vwretx in Cochrane (2008) and in Welch and Goyal (2008) during the period 1926-2004; we did the same for vwretd. The high correlation of $\approx 0.99$ convinced us to proceed with the extended analysis for the U.S. In the U.S. data from 1984-2018, we used the CPI from Shiller's website (2019), since these are an exact match of Cochrane's CPI from 1926-2004. We simply extended the data to include the period 2005-2018 from the website of Goyal (2020).

The Norwegian Stock Market data was collected from the Oslo Stock Exchange (2020). We received access to the monthly index levels and risk-free quarterly rates (accessed at Bernt (2020)) and calculated the market value of aggregated dividends for the Norwegian market. We must emphasize that the data is restricted to students and researchers at BI. As in Cochrane (2008), the data is annualized. We also retrieved the CPI directly from SSB (2020). We must emphasize that the stock market data is only available from 1984 to 2018. This restriction is due to the fact that the Norwegian index was not fully developed before 1984, which is why our comparative analysis with the U.S starts from 1984 (Oslo Stock Exchange (2020)). The proxy for risk-free rate in Norway comes from Bloomberg, where we accessed the three-month NIBOR rate (Bloomberg (2020)). However, this data was not available before 1986. We used the the nominal lending rate for banks in Norway to obtain the three-month NIBOR rate for 1984 and 1985 and regressed $NIBOR_{1986-2016} = \alpha + \beta[\text{Nominal lending rate}_{1986-2016}] + \varepsilon$, with $R^2$ close to 0.962. The obtained $\alpha \approx -0.022$ and $\beta \approx 1.009$ are used to calculate the three-month NIBOR rate for 1984 and 1985. This is similar to the method used by Welch and Goyal (2008) who regressed commercial papers as the proxy for the risk-free rate between 1871-1925.

In our Methodology, we regress the long horizon for Norwegian and U.S data in the same way that Cochrane did in his article. The lags used are 1, 3, 5, 7, and 11 years for the data from 1984-2018 for both markets. We have checked the maximum number of possible lags for the long-run forecast for our smallest dataset (the Norwegian data), and obtained a maximum lag of 11 years for the Norwegian market, and 20 for the U.S. If we were to use lags up to 20 years, as Cochrane (2020), we would need more data for the Norwegian market. Due to the structure of the long-run forecasts from Cochrane (2008), it requires three years of data observations in order for the model to forecast of one more lag. Therefore, to compare and analyse long-run regressions on equal premises, we have used maximum lag of 11 years for the forecast long-run forecasts.

## 5 Results and analysis

This section mainly provides the results and analysis of the tests we have described in the methodology and use these to compare the Norwegian and U.S. market data from 1984-2018. We present a comparison between the Norwegian- and the U.S market. We will illustrate two panels for each table, where the first panel will represent the Norwegian market data for 1984-2018, and the second panel will represent the U.S. market data for 1984-2018. This will be the main comparison in our analysis. Lastly, we present the complete replication of Cochrane's findings (2008), including the U.S data for 1926-2004, as well as an extension of the U.S. data for 1926-2018, which will include observed differences that should be noted for the U.S. data from 1926-2004/2018 relative to the U.S sample starting from 1984. Looking at these differences allows us to examine the changes in parameters over a longer horizon for the U.S. market. The extension and all tables from the U.S. extension is available in section extension and appendix.

and use these to compare the Norwegian and U.S. market data from 1984-2018. We present a comparison between the Norwegian- and the U.S market. We will illustrate two panels for each table, where the first panel will represent the Norwegian market data for 1984-2018 and the second panel will represent the U.S. market data for 1984-2018. This will be the main comparison in our analysis. Lastly, we present the complete replication of

### 5.1 Simple statistics

To look at forecast predictability, we will first of all look at plain statistics for the one-period regressions. The return has a standard deviation of 0.27, while the risk free rate has a standard deviation of 0.04 in the Norwegian data. In this setting, we consider the historical mean as our proxy for expected return and express risk as the data's standard deviation (volatility). The return of $S\&P$ 500 comes with a higher expected return than our proxy for the risk-free rate, although with a higher risk in terms of the standard deviation - which is expected. After all, higher return involves higher risk (Sharpe (1994)). All of the variables reject the joint normality test of Jarque-Bera (Thadewald and Büning (2007)), with test statistics exceeding the critical value of 5.99 with two degrees of freedom. Financial data tends to exhibit characteristics of leptokurtosis (i.e the distribution has fat tails and a higher mode) and a left-skewed distribution (Brooks (2014)). The negative skewness entails that high negative return is more likely than positive return of the same magnitude (Brooks (2014)). Yet again, these common features seem to be the case for return for both markets.

| | | **Correlation, std on diagonal** | | | | | **Other statistics** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r** | **dd** | **dp** | **RF** | **CPI** | **skew** | **kurt** | **jointly** | **mean** |
| **Norwegian 1984-2018** | **r** | 0.27 | 0.21 | 0.18 | -0.17 | -0.10 | -1.09 | 2.33 | 7.56 | 0.11 |
| | **dd** | 0.21 | 0.26 | -0.37 | -0.10 | -0.30 | -0.95 | 2.12 | 6.39 | 0.09 |
| | **dp** | 0.18 | -0.37 | 0.39 | -0.69 | -0.26 | 0.42 | -0.18 | 15.79 | -3.56 |
| | **RF** | -0.17 | -0.10 | -0.69 | 0.04 | 0.77 | 0.85 | -0.39 | 20.93 | 0.06 |
| | **CPI** | -0.10 | -0.30 | -0.26 | 0.77 | 0.02 | 1.78 | 3.76 | 19.36 | 0.03 |
| **U.S 1984-2018** | **r** | 0.16 | 0.67 | 0.38 | 0.21 | 0.07 | -1.27 | 2.25 | 10.28 | 0.08 |
| | **dd** | 0.67 | 0.13 | 0.09 | -0.09 | -0.03 | 0.14 | 0.25 | 11.13 | 0.03 |
| | **dp** | 0.38 | 0.09 | 0.35 | 0.49 | 0.41 | 0.12 | -0.38 | 16.79 | -3.75 |
| | **RF** | 0.21 | -0.09 | 0.49 | 0.03 | 0.66 | 0.16 | -1.27 | 26.77 | 0.03 |
| | **CPI** | 0.07 | -0.03 | 0.41 | 0.66 | 0.01 | 0.51 | 0.65 | 9.60 | 0.03 |

Table 1: Statistics

r is log return at time t+1 and dd is log dividend change t+1, deflated by CPI. Dp is log dividend price ratio at time t, RF is T-bill Three months at t+1 and CPI is the consumption price index at t+1. In "Correlation, std on diagonal", the diagonal is standard deviation of the corresponding letters, and the rest is the correlation.

When the correlation is different from zero, the variation on one variable can cause some variation on other variables. Assessing the correlations, we can see that the correlation between return ($r_{t+1}$) and dividend price ratio ($dp_t$) in the Norwegian data is 0.18 and correlation between dividend growth ($dd_{t+1}$)

and ($dp_t$) is -0.37. Statistically, since both correlations are different from zero, it seems like dp has the potential to explain some of the variation on return and dividend growth. However, the U.S. market shows different correlations, where return shows high positive correlation with dividend-yield, while dividend-growth is close to zero correlation. This implies that return moves more independent to dividend-growth than the dividend-yield. Hence, more correlation in the independent variable and almost none in delta of the independent variable. Usually, we would expect the correlation between the risk-free rate and CPI (inflation) to be negatively correlated. A higher interest rate should lead to a higher yield to maturity on bonds, which is the benchmark for the risk-free rate (Folger (2016)). However, the effects of an increase or a decrease in interest rates usually propagates in the economy 2-4 years later (Folger (2016)). Hence, looking at the one-period correlation between CPI and the risk-free rate does not necessarily give much insight in this table.

### 5.2 Simple regressions and the VAR-system

To check for forecastability, we regressed return, excess return, and dividend growth on the dividend-price ratio in table 2 below. We also tested for statistical significance for the independent variables. The estimates of $\beta_r$ & $\beta_d$ are based on the non-forecastable null-hypothesis for return. Meaning, $\beta$ is calculated under a non-forecastable return hypothesis while dividend-growth is forecastable. Recall that the hypothesis written in methodology is as follows:

$$H_0 : b_r = 0, b_d = \rho\phi - 1$$

$$H_A : b_r = 1 - \rho\phi, b_d = 0$$

$H_0$ assumes that return is **unforecastable** and dividend-growth is **forecastable**. Alternative hypothesis assumes the opposite.

Looking at the Norwegian data, the dividend yield variable on dividend growth, $\beta_D$, presents itself as a significant explanatory variable with a t-stat of $(-2.69)$, exceeding the critical value of +-1.96 at the 5% level. The dividend-yield on return, $\beta_R$, is not statistically significant alone with a t-stat of 1.19. This is different from U.S. between 1984-2018 with a significant $\beta_R$, where t-stat is 2.42 and thereby exceeds the critical value. This regression alone does not provide considerable insight. We see some inconsistencies in terms of which variables show statistical significance for the U.S. compared to the Norwegian stock market. Some of the output for the U.S. has completely different implications than the output for the Norwegian market. Regarding dividend growth, however, the Norwegian data has a profoundly negative $\beta_D$ of $-8.474$, compared to the U.S $\beta_D$, which is close to zero and slightly positive. From what we refer to as *the economic intuition* in this thesis, high dividend-yield gives low prices, leading to lower return and a decrease in future expected dividend growth (Cochrane (2008)). Hence, the correlation between dividend growth and dividend yield should be negative. This is why we find the negative $B_d$ in by the Norwegian data to have a high degree of economic significance. The fact that the statistics look so different for the U.S. market is something we find very significant and will discuss further later on.

From the statistics above, we obtain the correlation for the Norwegian data of $(corr[\Delta d_{t+1}, dp_t] < 0) \approx -0.3656$ and $(corr[\Delta d_{t+1}, dp_t] > 0) \approx 0.0929$ for the U.S. data. Recalling *the economic intuition*, we would expect a negative relationship between the current dp-ratio and the future dividend growth. We can say that this statistically holds for the Norwegian data. However, it does not hold for the U.S. data (which has a positive correlation instead). Looking at the $R^2$, we see that the return comes with higher $R^2$ in the U.S. data compared to the Norwegian data, while it is lower in the U.S data for dividend growth. This

| | Regression | $\beta$ | se | $t(\beta)$ | $R^2$ | Stdx*b |
|---|---|---|---|---|---|---|
| Norwegian 1984-2018 | $R_{t+1} = \alpha_R + \beta_R(D_t/P_t) + \varepsilon^R_{t+1}$ | 2.904 | 2.446 | 1.19 | 0.0197 | 0.039 |
| | $R_{t+1} - R^f_t = \alpha_R + \beta_{R-RF}(D_t/P_t) + \varepsilon^R_{t+1}$ | 4.592 | 2.424 | 1.89 | 0.0470 | 0.061 |
| | $D_{t+1}/D_t = \alpha_D + \beta_D(D_t/P_t) + \varepsilon^D_{t+1}$ | -8.474 | 3.149 | -2.69 | 0.1813 | 0.113 |
| | $r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$ | 0.122 | 0.079 | 1.55 | 0.0331 | 0.048 |
| | $\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1}$ | -0.238 | 0.147 | -1.62 | 0.1337 | 0.094 |
| U.S 1984-2018 | $R_{t+1} = \alpha_R + \beta_R(D_t/P_t) + \varepsilon^R_{t+1}$ | 6.157 | 2.547 | 2.42 | 0.1166 | 0.056 |
| | $R_{t+1} - R^f_t = \alpha_R + \beta_{R-RF}(D_t/P_t) + \varepsilon^R_{t+1}$ | 5.226 | 2.533 | 2.06 | 0.0878 | 0.048 |
| | $D_{t+1}/D_t = \alpha_D + \beta_D(D_t/P_t) + \varepsilon^D_{t+1}$ | 0.559 | 2.390 | 0.23 | 0.0013 | 0.005 |
| | $r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$ | 0.177 | 0.067 | 2.64 | 0.1461 | 0.063 |
| | $\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1}$ | 0.035 | 0.063 | 0.56 | 0.0086 | 0.012 |

Table 2: Forecasting

Capital letters are real returns using CPI, and small letters are logs of corresponding letters. $R_{t+1}$ is return, $\frac{D_{t+1}}{D_t}$ is real dividend growth, and $\frac{D_t}{P_t}$ is dividend price ratio. $R^f_t$ is the risk-free rate at time t. $b_r$ is coefficient for log return and $b_d$ is coefficient for dividend growth. The coefficient for return is estimated under the null that assumes that return is not forecastable *while* dividend growth is; The opposite applies for dividend-growth coefficient

implies better explanatory power for return for the U.S. and better explanatory power for dividend-growth for the Norwegian market. Moving on, we look at Cochrane's VAR system (2008).

By using the VAR-system developed by Cochrane, we obtained estimates for each coefficient and obtained identity (5), $b_r = 1 - \rho\phi + b_d$; all of which is described in the methodology.

Table 3 shows the estimated $\hat{\beta}_r$, $\hat{\beta}_d$, $\hat{\beta}_{dp}$ and the correlation between the shocks from dependent variables. We obtained a negative correlation between shocks in return and dividend yield for both markets: $Corr(\epsilon_r, \epsilon_{dp}) \approx -0.642$ for the Norwegian market and $Corr(\epsilon_r, \epsilon_{dp}) \approx -0.527$ for the U.S. This suggests that a shock increase in price should lead to an increase in return and lower the dividend yield (Linearization), as emphasized in the article of Cochrane (2008). Recalling the *economic intuition*, we can immediately say that we find the value for this specific correlation meaningful for both markets, due to the relation-

| | | Estimates | | | $\varepsilon$ s. d. diagonal and corr | | | Null 1 | Null 2 |
|---|---|---|---|---|---|---|---|---|---|
| | | $\hat{\beta}$ | $\sigma(\hat{\beta})$ | Implied | r | Δd | dp | $\beta, \Phi$ | $\beta, \Phi$ |
| **Norwegian 1984-2018** | **r** | 0.122 | 0.079 | 0.122 | 0.261 | 0.302 | -0.642 | 0.000 | 0.000 |
| | **Δd** | -0.238 | 0.147 | -0.237 | 0.302 | 0.239 | 0.538 | -0.360 | -0.038 |
| | **dp** | 0.659 | 0.156 | 0.658 | -0.642 | 0.538 | 0.306 | 0.659 | 0.990 |
| **U.S 1984-2018** | **r** | 0.177 | 0.067 | 0.177 | 0.151 | 0.687 | -0.527 | 0.000 | 0.000 |
| | **Δd** | 0.035 | 0.063 | 0.036 | 0.687 | 0.133 | 0.255 | -0.141 | -0.033 |
| | **dp** | 0.879 | 0.058 | 0.878 | -0.527 | 0.255 | 0.116 | 0.879 | 0.990 |

Table 3: The Vector Autoregression (VAR)-system

Each row represents the one-period regression from the Vector Autoregression (VAR)-system described in the methodology. For instance, the first row uses the regression $r_{t+1} = \alpha + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$, equivalent to identity 1; the same applies for the second and third row for each set. $r$, $\Delta d$ and $dp$ are in time t+1. The implied values are calculated by solving the corresponding dependent variable in identity 5. $\rho$ is defined as the constant Taylor-approximated point estimate for the dividend yield and is used to calculate the implied value for the corresponding dependent variable in table 3 above. The null columns are the coefficients that are used in the simulations under the null.

ship between return and dividend yield. The same applies for the correlation between dividend growth and dividend yield This is why we believe the correlation between dividend growth and dividend yield should be positive, which is exactly what was revealed for both markets.

By looking at the regression for the dividend-growth (identity 3), rewriting and solve for $\phi$, we obtain $b_r = 1 - \rho\phi + b_d \rightarrow \frac{b_d - b_r + 1}{\rho} = \phi$. The negative $corr_N[\Delta d_{t+1}, dp_t]$ and positive $corr_N[r_{t+1}, dp_t]$ must, by construction, lead to a low $\phi$ when the constant level of $\rho$ is high. For the U.S data, this means that positive $corr_{U.S}[\Delta d_{t+1}, dp_t]$ and $corr_{U.S}[r_{t+1}, dp_t]$, and high $\rho$, leads to high $\phi$. In short, we believe that the difference in correlation between $d_{t+1}$ and $dp_t$ for the corresponding markets is what makes the difference in the autocorrelation of dividend-yield for $\phi_{Norway}$ and $\phi_{U.S}$.

### 5.3 The importance of phi

In the previous section, we looked at the Vector Autoregression (VAR)-system, identity (1)-(3), and found that $\phi$ is different in the Norwegian data compared to the U.S. Data. The dividend growth was also different due to the negative correlation between $\Delta d_{t+1}$ and $dp_t$ in the Norwegian data, compared to the positive correlation in the U.S. data. From the *economic intuition*, the autocorrelation function for dividend-yield $\phi$ is expected to be negatively correlated with return. We are interested in examining how the effects of an increase in $\phi$ play out on the coefficients for real and excess return in each of the datasets.

| | | Percent probability values | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Real returns | | | | Excess returns | | | |
| | $\phi$ | $\beta_r$ | $\phi$ | $br^{lr}_{min}$ | $br^{lr}_{max}$ | $\beta_r$ | $\phi$ | $br^{lr}_{min}$ | $br^{lr}_{max}$ |
| **Norwegian 1984-2018** | **0.640** | 26.07 | 4.98 | 14.24 | 14.33 | 14.00 | 1.33 | 3.34 | 3.38 |
| | **0.659** | 25.99 | 6.99 | 15.24 | 15.38 | 13.82 | 2.04 | 3.65 | 3.69 |
| | **0.700** | 25.64 | 12.79 | 17.64 | 17.76 | 13.58 | 3.95 | 5.03 | 5.10 |
| | **0.800** | 25.38 | 42.08 | 25.63 | 25.77 | 13.05 | 18.69 | 10.15 | 10.24 |
| | **0.900** | 24.87 | 78.46 | 38.26 | 38.40 | 11.82 | 54.63 | 20.12 | 20.23 |
| | **0.960** | 25.22 | 90.75 | 50.58 | 50.70 | 11.27 | 76.40 | 30.98 | 31.11 |
| | **0.980** | 25.51 | 93.38 | 55.80 | 55.95 | 11.54 | 81.65 | 36.33 | 36.45 |
| | **0.990** | 25.55 | 94.34 | 58.74 | 58.84 | 11.58 | 84.32 | 40.07 | 40.18 |
| | **1.000** | 25.89 | 95.01 | 61.80 | 61.91 | 11.63 | 86.25 | 42.78 | 42.92 |
| | **1.010** | 24.49 | 95.92 | 64.70 | 64.79 | 10.56 | 88.73 | 45.89 | 46.01 |
| | **Draw** | 25.34 | 18.57 | 15.33 | 15.41 | 13.60 | 15.30 | 8.18 | 8.19 |
| **U.S 1984-2018** | **0.840** | 23.83 | 3.47 | 2.15 | 2.21 | 26.19 | 4.43 | 3.59 | 3.66 |
| | **0.879** | 23.59 | 4.72 | 3.27 | 3.34 | 26.22 | 6.50 | 5.55 | 5.69 |
| | **0.900** | 23.03 | 5.92 | 4.43 | 4.54 | 25.78 | 7.64 | 6.76 | 6.86 |
| | **0.960** | 22.65 | 10.27 | 9.76 | 9.94 | 25.49 | 13.48 | 13.23 | 13.39 |
| | **0.980** | 22.69 | 12.11 | 12.49 | 12.69 | 25.54 | 16.43 | 16.55 | 16.76 |
| | **0.990** | 23.02 | 13.91 | 14.63 | 14.81 | 25.46 | 17.95 | 18.43 | 18.67 |
| | **1.000** | 22.88 | 15.39 | 16.56 | 16.75 | 25.62 | 19.70 | 20.17 | 20.37 |
| | **1.010** | 21.58 | 16.74 | 19.13 | 19.35 | 24.27 | 22.12 | 22.98 | 23.19 |
| | **Draw** | 24.34 | 5.73 | 5.26 | 5.36 | 26.75 | 7.38 | 7.02 | 7.14 |

Table 4: Increasing $\phi$ (phi)

Table 4 shows the probability values that the simulated coefficients are larger than the corresponding coefficients for different values of $\phi$. The probability value 25.99 for $b_r$ is calculated as $\frac{\sum b_r^{sim} > b_r^{data}}{50.000}$ using $\phi = 0.659$ in the simulations. The coefficient for return is estimated under the unforecastable return-null *while* dividend growth is forecastable. The coefficient for dividend-growth is estimated under the null that assumes that dividend-growth cannot be forecast *while* return is forecastable.

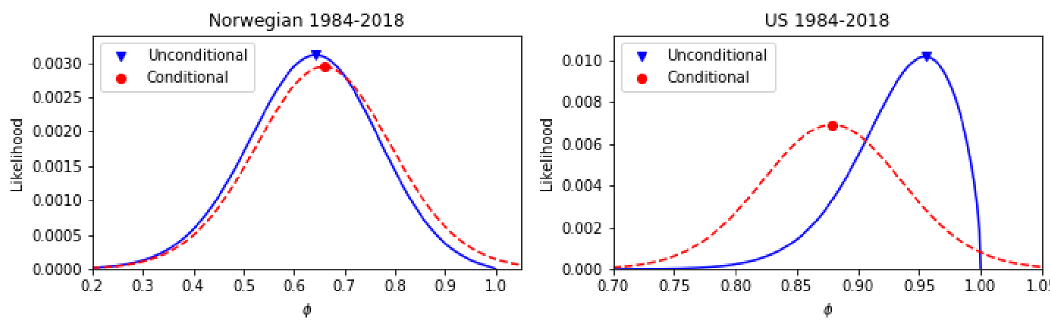



Figure 1: Conditional and unconditional likelihood

The figures show unconditional and conditional maximum likelihood function where the $\phi$ is peaked at around 0.659. The figure is an Autoregression (AR)[1] process of dividend yield, where the conditional likelihood uses the first data point as fixed. The unconditional likelihood adds the log probability value of the first data point. The Likelihood function for $\phi$, autoregressive coefficient for dividend yield $d_{t+1} - p_{t+1} = \alpha_{t+1} + \beta_{dp}(d_t - p_t) + \varepsilon^{dp}_{t+1}$.

Here are the density functions for $\phi_{Norway}$ and $\phi_{U.S}$. We can immediately see that the distributions for $\phi$ are a bit different. Firstly, the distribution of the unconditional mean $\phi_{U.S}$ is more left skewed than the $\phi_{Norway}$. The density function of $\phi_{Norway}$ is closer to a normal distribution, both for the conditional and unconditional mean. When the mean is conditional in this case, it depends on some past information, from its previous value, which in this case involves having the first data point as fixed (Cochrane (2008)). Therefore, we are not surprised to see that the shape of the conditional curves are quite similar. In terms of the range $\phi$, the spread of $\phi_{Norway}$ is wider than $\phi_{U.S.}$. In fact, the distribution is so different that, if one were to present the two curves on the same graph, the end of the density function for $\phi_N orway$ would only *just* enter the interval at which we find the mode for $\phi_{U.S.}$. Due to the fact that the value of $\phi$ is based on the mode (the peak value of the distribution) of the respective distributions, this affects the way our difference in results for the two markets. We saw in table 3 how $\rho$ and $corr_N[\Delta d_{t+1}, dp_t]$ and $corr_N[r_{t+1}, dp_t]$ affects the value of $\phi$. However, we will later discuss how the lower $\phi$ for Norway and a higher $\phi$ for the U.S. can affect their respective long-run regressions.

The $\phi$ in Identity (3) represents the autocorrelation on dividend yield. By using the bootstrapping method, we used the mode of distribution for both $\phi$ and $\rho$ as the value of the respective coefficients, which is described in the appendix. We obtained the autocorrelation value $\phi_{Norway}$ = 0.659 for the Norwegian data and $\phi_{U.S}$ = 0.879 for the U.S data, denoted with corresponding subscripts for the two markets. The probability values for the $b_r$ exceeding its sample value stay around 22-24% for any value of $\phi$. As $\phi$ increases, $b_d$ increases, but does not explode until $\phi$ reaches around 0.9 in the Norwegian data. From $\phi = 0.9$ through $\phi = 1.01$, there are no massive changes. Excess returns are better in the sense that the probability values are lower, which insinuates less random estimates with respect to the forecast. At $\phi = 1$, the probability value is close to 96%.

In terms of the difference in correlation, which we previously discussed, we believe that this is the potential root of the difference in $\phi$. Also, we do not find it particularly comforting that bd more or less explodes and becomes pretty much unstable due to the density function of the Norwegian data being narrower than the U.S data. In other words, the $\phi_{U.S.}$ is more sensitive to changes without making significant changes in estimates than $\phi_{Norway}$. This lack of stability for the U.S. is more susceptible to spurious inference if $\phi$ increases too much. We saw earlier in table 2 that the intuition of the sign of $B_d$ for the Norwegian market seems reasonable with regard to *economic intuition*. This is aligned with the defined identities, and thereby with Campbell and Shiller (1988).

Now that we have looked at the regressions, the Vector Autoregression (VAR)-system, we are interested in looking at the probabilities for the predicting variables in the regressions. The probabilities will indicate the statistical likelihood to reject the unforeseeable null-hypotheses and thereby enable us to say how likely it is for an estimate to exceed the sample value by pure chance.

### 5.4 Short-term forecasting

Table 5 shows the probabilities of the instances where the coefficients are greater than their respective estimated sample values under the null-hypothesis. The bootstrapping method enables us to visualize the (lack of) persistence of dividend yield (Verdickt, Annaert, and Deloof (2019)). The return coefficient has a probability of 25.812%, which is the probability that the coefficient $b_r \approx 0.122$ is random. The probability for $b_d$ is much lower, 7.094% ,which implies that dividend growth yields more fitting properties for forecasting by, in principle, being less random over time. The excess return gives far better probabilities than the real values for Norwegian data. The U.S. data from 1984-2018 gives almost the same probability for return, but much lower for the t-stat than the Norwegian probabilities. These differences, as well as the differences in the probabilities for $b_d$ and $t_d$, are worth taking notes of.

| | | $\beta_r$ | $t_r$ | $\phi$ | $t_d$ |
|---|---|---|---|---|---|
| **Norwegian 1984-2018** | **Real** | 25.812 | 13.752 | 7.094 | 1.834 |
| | **Excess** | 14.200 | 4.444 | 1.986 | 0.672 |
| **U.S 1984-2018** | **Real** | 23.454 | 3.418 | 4.968 | 2.924 |
| | **Excess** | 25.800 | 4.878 | 6.204 | 5.050 |

Table 5: Monte Carlo Simulation by Bootstrapping

The columns give the probability that the coefficients are larger than the corresponding coefficients in the sample data. Example is $b_r^{real} = \frac{\sum b_r{}^{sim(real)} > b_r^{data(real)}}{50.000}$. 50.000 samples of $b_r$, $t_r$, $b_d$ and $t_d$ in a Monte Carlo simulation and calculated how many of these samples that were greater than the sample coefficients. The coefficients are estimated under the unforecastale return-null *while* dividend growth is forecastable.

Based on our Norwegian data results, the t-stat from the simulation is greater than the historical mean more often on real return than excess return. However, this is not true for the U.S. market. Although the $b_r$ coefficient implies less accuracy on the returns by using this method, the other parameters deviate in relatively few cases - especially we look at the probability values for the excess return. Looking at the coefficients from Table 2, none of the log parameters for r and dd are significant at 5% level in the Norwegian data. However, the probabilities for return are higher than the probabilities for dividend growth, assuming a null hypothesis that return is unpredictable. It seems like the U.S performs better than the Norwegian market for return for a joint hypothesis that return is unforecastable. This is the same as what Engsted and Pedersen (2010) found. Therefore, it is easier to reject dividend growth forecastability than return unforecastability for both markets. Moreover, the one-period regression is one side of our research. Although we are interested in the regression to find indications for the one period, an equally significant aspect of this research is to test the model's long-run forecasting for the two markets.

### 5.5 How does this work in the long-run?

Using the Cochrane System (2008), as applied to Campbell and Shiller's model (1988), we calculated the long-run coefficients for beta, standard error and t-stat for both markets. Calculations are available in Appendix D.3 long-run.

| | | $\beta^{lr}$ | se | t | % prob. values |
|---|---|---|---|---|---|
| Norwegian 1984-2018 | $r_{t+1}$ | 0.34 | 0.26 | 1.33 | 15.24-15.35 |
| | $\Delta d_{t+1}$ | -0.66 | 0.26 | 1.33 | 15.24-15.35 |
| | Excess $r_{t+1}$ | 0.51 | 0.25 | 2.04 | 3.98-4.04 |
| U.S 1984-2018 | $r_{t+1}$ | 1.25 | 0.49 | 2.55 | 3.30-3.40 |
| | $\Delta d_{t+1}$ | 0.25 | 0.49 | 2.55 | 3.30-3.40 |
| | Excess $r_{t+1}$ | 1.11 | 0.48 | 2.30 | 5.32-5.44 |

Table 6: Long-run horizon

Table 6 shows the long-run coefficients. Note that the "lr" superscript denotes long-run coefficients. $\hat{\beta}_r^{lr}$ is calculated as $\hat{\beta}_r^{lr} = \frac{\hat{b_r}}{1-\rho\hat{\phi}}$ where $\hat{b_r}$ is the coefficient from the regression $r_{t+1} = \alpha + \hat{\beta}_r(d_t - p_t) + \varepsilon^r_{t+1}$. $\hat{\phi}$ comes from the autocorrelation $d_{t+1} - p_{t+1} = \alpha + \hat{\phi}_{dp}(d_t - p_t) + \varepsilon^{dp}_{t+1}$. The % probability values are the probability range under the null-hypothesis $\phi$ from identity (10) in Methodology. The test-statistic for $b_d^{lr}$ is calculated under the assumption that $\hat{b}_d^{lr} = -1$, while the null-hypothesis for $\hat{b}_d^{lr}$ assumes $\hat{b}_r^{lr} = 0$.

The point estimates for the Norwegian long-run return coefficient $b_r^{lr}$= 0.34 and long-run dividend growth coefficient $b_d^{lr} = -0.66$ signal that about 66% dividend yield volatility comes from dividend growth and about 34% comes from log returns. The negative correlation between dividend growth and dividend-yield implies that the long-run $b_d^{lr}$ should also be negative. This seems to be the case for the VAR system's dividend growth regression shown in section 5.2 of this thesis and the long-run forecast for the Norwegian data in the table above. Another interpretation is that the U.S. data shows considerably higher significance for all long-run variables, compared to the Norwegian data. The forecastability of the return is also consistent with the findings of Campbell and

Shiller (1988), who argue that return can, to some extent, be forecastable. However, similar to what we saw in table 1, the long-run dividend growth estimate imply the opposite of what we would expect for the U.S by having, what we would consider, the "wrong" sign. It returns a point estimate of 0.25. This is something Cochrane (2008) himself regarded as a misleading result.

The range of simulated probability values of return for the Norwegian data of $\approx 15.24 - 15.35\%$ reflects the range of probabilities that the simulated coefficients are larger than the sample data coefficients. These are higher than the probability values in the U.S data, $b_r^{lr}(sim) > b_r^{lr}(data)$. The unanticipated positive $corr[\Delta d_{t+1}, dp_t] \approx 0.0929$ and the positive sign of the $\hat{\beta_d}$ in the U.S data, might affect these probability values and the t-stats. In the U.S data, more than 125% of the dividend yield volatility comes from the return, $b_r^{lr}$, and about 25% comes from dividend growth, $b_d^{lr}$. Cochrane (2008) also notes that such a decomposition can return values above 100% and less than 0% when not orthogonalized. This implies that most of the volatility comes from return and significantly less comes from the dividend for the U.S., something which appears to be connected to the difference in correlation. Despite these probability values, it is interesting to look at different values of $\phi$. Phrased differently, high $b_r$ should lead to low *phi* due to the correlation between return shocks and dp shocks, being largely negative. Similarly, the correlation between dividend-growth shocks and shocks in the dividend-yield is highly positive, meaning that a large $\phi$ should produce high $b_d$. Further, in the Norwegian data, the probability values for return decreased from short- to long-run under the null-hypothesis for unforecastable return (identity 10), while it decreased significantly much more in the U.S. market. This can be interpreted as the return unforecastability more likely to be rejected under the null for the U.S. market in the long-run, compared to the Norwegian market. From this, we find stronger evidence against the null for unforecastable return than the forecastable dividend growth for the U.S. market compared to the Norwegian. Hence, there is evidence that return forecastability improves in the long-run only for the U.S. market, which is consistent with the findings of Engsted and Pedersen (2010) and Monteiro (2018) who found that that return is more forecastable in the U.S. than the Norwegian market.

From the methodology, the variation in dp ratio is defined as *either* stemming from the covariance between return and dp, or dd and dp. Cochrane (2008) stressed the importance of looking at *how much* of each component's variation moves prices the most, instead of primarily focusing on *if* one of the two moves prices the most. We determine the impact of variation by looking at the correlations between the variables, which was shown in in table 3. The higher the absolute value of the correlation, the greater the impact on the independent variable. Hence, most of the variation comes from return in the U.S data, while most of the variation comes from dividend-growth for the Norwegian market. In sum, the long-run estimates for the Norwegian data are different from estimates for the U.S data.

## 5.6 Long-run forecasting

After looking at the long-run regression, we are interested in looking at the long-run forecasting in the same way as we did for one period in the section forecasting. This section also focuses on the weighted and unweighted long-run regressions, where r is included in the weighted formula. We assess the regression based on direct estimates and implied estimates and different lags. The implied long-run return coefficient was calculated by solving for $Br^{lr}$ in identity 5. The value of the direct estimates descends from the formula at the top of Table 7.

| | | Weighted $\sum_{j=1}^{k}\rho^{j-1}r_{t+j} = \alpha + \beta_r^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | | Unweighted $\sum_{j=1}^{k}r_{t+j} = \alpha + \beta_r^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Direct | | | | Implied | | | Direct | | | | Implied | | |
| | k | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ |
| **Norwegian 1984-2018** | 1.0 | 0.12 | 25.86 | 25.79 | 0.03 | 0.12 | 25.86 | 25.79 | 0.12 | 25.86 | 25.79 | 0.03 | 0.12 | 25.86 | 25.79 |
| | 3.0 | 0.11 | 51.12 | 63.78 | 0.01 | 0.25 | 20.69 | 32.68 | 0.10 | 51.96 | 64.94 | 0.01 | 0.26 | 20.51 | 32.86 |
| | 5.0 | 0.00 | 68.06 | 82.47 | 0.00 | 0.30 | 17.66 | 38.86 | -0.01 | 69.52 | 83.91 | 0.00 | 0.31 | 17.26 | 39.30 |
| | 7.0 | -0.36 | 89.58 | 96.54 | 0.11 | 0.33 | 16.32 | 43.70 | -0.44 | 91.11 | 97.25 | 0.14 | 0.34 | 15.77 | 44.65 |
| | 11.0 | -0.32 | 88.70 | 93.49 | 0.04 | 0.34 | 15.48 | 50.04 | -0.37 | 88.62 | 93.08 | 0.04 | 0.36 | 14.73 | 51.88 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | 0.34 | 15.27 | 58.96 | ∞ | ∞ | ∞ | ∞ | 0.36 | 14.51 | 63.66 |
| **U.S 1984-2018** | 1.0 | 0.18 | 23.41 | 22.71 | 0.15 | 0.18 | 23.41 | 22.71 | 0.18 | 23.41 | 22.71 | 0.15 | 0.18 | 23.41 | 22.71 |
| | 3.0 | 0.41 | 28.41 | 28.64 | 0.28 | 0.46 | 17.56 | 20.01 | 0.42 | 28.48 | 28.67 | 0.28 | 0.47 | 17.43 | 19.95 |
| | 5.0 | 0.62 | 28.42 | 29.67 | 0.41 | 0.67 | 12.77 | 17.42 | 0.64 | 28.47 | 29.64 | 0.40 | 0.70 | 12.44 | 17.24 |
| | 7.0 | 0.78 | 28.25 | 29.98 | 0.63 | 0.82 | 9.19 | 15.33 | 0.83 | 28.21 | 29.89 | 0.62 | 0.87 | 8.65 | 14.97 |
| | 11.0 | 1.05 | 26.05 | 27.60 | 0.85 | 1.02 | 5.62 | 12.90 | 1.15 | 26.05 | 27.47 | 0.81 | 1.11 | 4.91 | 12.32 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | 1.25 | 3.23 | 14.31 | ∞ | ∞ | ∞ | ∞ | 1.46 | 2.58 | 15.19 |

Table 7: Long-run forecasting- Return

Table 7 shows the long-run coefficients for the regression $b_r^{(k)} = \sum_{j=1}^{k}\rho^{j-1}\phi^{j-1}b_r$, where k is the number of lags. The Table includes direct and indirect estimates for weighted and unweighted regression, respectively (calculation available in direct and indirect estimates for weighted and unweighted regression). The $\hat{b}_r^{(k)}$ is the coefficient of the long-run estimate for the corresponding lag. The *phi_data* is the probability that the simulated long-run coefficient $\hat{b}_r^{(k)}$ exceeds the sample value $\hat{b}_r^{(k)}$, for each data set, under the assumption of null hypothesis for unforecastable return.

The long-run coefficient $\hat{b}_r^{(k)}$ declines in the Norwegian data as we increase the number of lags, k. The probability values increase substantially. This is due to the negative correlation between the return coefficient and the log of dp-ratio, when $j > 3$, which is, in comparison, substantially more negative in the Norwegian market. Historically speaking, the aggregated stock market returns tends to increase in the long-run (Wohlner (2020)). Because of this, we would expect the long-run forecast to increase in value as k increases (Maio and Santa-Clara (2015) and Cochrane (2008)); similar to what we see for the U.S. in our sample, and for the extended periods. Surprisingly, the Norwegian market shows a pattern for the return in the long-run, which is the opposite to what was anticipated. This is unexpected in terms of what the model itself should forecast, considering that all variables were the same. However, the articles by Monteiro (2018) and Engsted and Pedersen (2010) show the same results and emphasize that this seems to be the case when using the dividend-yield model in most countries except for the U.S. Moreover, the dividend-yield autocorrelation $\phi$ is relatively low, making it harder to generate a positive coefficient for return in the long-run (Cochrane (2008)). In other words, the lower value of $\phi_{Norway}$ causes the long-run coefficient for return to decrease over time. This is exactly what Fama and French (1988) concluded; "high autocorrelation causes the variance of expected returns to grow faster than the return horizon" (Fama and French (1988, p. 1)).

Based on the declining p-values for the implied estimates, the expectation would be that it would forecast better and better - in theory. By interpreting the methodology's hypotheses, we see that an increase in lags gives less evidence against the unforecastable null for return in the Norwegian data. The U.S. data provides better evidence against the null, but it is still worse than evidence for the one period. That means that the probabilities are not giving better forecastability in the long-run compared to the direct estimates presented by Fama and French (1988). The absence of forecastability and reduced evidence against the null-hypothesis for return (identity 10) might be due to the reduced data sample size or due to the 2008 financial crisis. The long-run probability values in Cochrane's sample during 1926-2004 show a decline. When we extend the data sample to include the years up to 2018 (shown in section 5.8: Extension and the appendix), the probability values remains almost unchanged in the long-run. Therefore, this indicate that forecastable estimation is improved when using longer samples. Both weighted and unweighted estimates barely differ in value, which is expected due to $\rho \approx 1$. Moreover, the $R^2$ increases substantially for the U.S. market, but stays close to zero for the Norwegian market. This supports that return is more forecastable in the U.S. than Norway, which is same as Fama and French (1988).

An equally important aspect of this research is to examine the unforecastable null for dividend growth. This is an extension of Cochrane (2008) and is therefore not included in his paper.

| | | **Weighted** $\sum_{j=1}^{k}\rho^{j-1}\Delta d_{t+j} = \alpha + \beta_d^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | | **Unweighted** $\sum_{j=1}^{k}\Delta_{t+j} = \alpha + \beta_d^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Direct** | | | | **Implied** | | | **Direct** | | | | **Implied** | | |
| | **k** | $\hat{\beta}_d^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_d^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $\hat{\beta}_d^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_d^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ |
| **Norwegian 1984-2018** | 1.0 | -0.24 | 27.66 | 33.22 | 0.13 | -0.24 | 27.66 | 33.22 | -0.24 | 27.66 | 33.22 | 0.13 | -0.24 | 12.58 | 0.42 |
| | 3.0 | -0.31 | 48.74 | 66.05 | 0.10 | -0.49 | 20.92 | 42.48 | -0.32 | 49.16 | 66.48 | 0.10 | -0.50 | 20.72 | 42.72 |
| | 5.0 | -0.71 | 35.93 | 51.18 | 0.34 | -0.59 | 16.34 | 50.76 | -0.76 | 35.16 | 50.29 | 0.35 | -0.62 | 15.72 | 51.57 |
| | 7.0 | -0.95 | 33.20 | 49.68 | 0.55 | -0.63 | 13.96 | 57.43 | -1.04 | 31.98 | 48.24 | 0.56 | -0.67 | 13.10 | 58.79 |
| | 11.0 | -1.09 | 40.17 | 55.68 | 0.39 | -0.66 | 12.58 | 66.00 | -1.20 | 40.03 | 54.21 | 0.37 | -0.69 | 11.20 | 68.39 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | -0.66 | 12.16 | 78.80 | ∞ | ∞ | ∞ | ∞ | -0.70 | 10.81 | 85.99 |
| **U.S 1984-2018** | 1.0 | 0.04 | 82.18 | 92.66 | 0.01 | 0.04 | 82.18 | 92.66 | 0.04 | 82.18 | 92.66 | 0.01 | 0.04 | 2.29 | 0.34 |
| | 3.0 | -0.01 | 76.34 | 86.90 | 0.00 | 0.09 | 82.31 | 92.38 | -0.01 | 76.24 | 86.74 | 0.00 | 0.09 | 82.33 | 92.38 |
| | 5.0 | -0.07 | 76.04 | 84.41 | 0.01 | 0.13 | 82.52 | 92.13 | -0.07 | 75.72 | 84.08 | 0.02 | 0.16 | 82.54 | 92.12 |
| | 7.0 | -0.14 | 75.63 | 82.63 | 0.08 | 0.16 | 82.66 | 91.86 | -0.15 | 75.14 | 81.99 | 0.09 | 0.19 | 82.68 | 91.82 |
| | 11.0 | -0.18 | 78.01 | 82.18 | 0.15 | 0.20 | 82.86 | 91.51 | -0.21 | 76.61 | 80.86 | 0.15 | 0.24 | 82.88 | 91.46 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | 0.25 | 83.17 | 91.34 | ∞ | ∞ | ∞ | ∞ | 0.29 | 83.40 | 91.82 |

Table 8: Long-run forecasting power- Dividend

Table 8 shows the long-run coefficients for the regression $b_d^{(k)} = \sum_{j=1}^{k}\rho^{j-1}\phi^{j-1}b_d$, where k is the number of lags. The Table includes direct and indirect estimates for weighted and unweighted regression, respectively (calculation available in direct and indirect estimates for weighted and unweighted regression). The $\hat{b}_d^{(k)}$ is the coefficient of the long-run estimate for the corresponding lag. The $\phi_{Data}$ is the probability that the simulated long-run coefficient $\hat{b}_d^{(k)}$ is lower than sample value $\hat{b}_d^{(k)}$, for each data set, under the assumption of null hypothesis for unforecastable dividend growth.

Given a null-hypothesis assuming unforecastable dividend-growth, we see similarities in explanatory power for long-run regression for both markets for 1984-2018, although considerably higher for the Norwegian market. Moreover, we obtain decreasing probability values for both markets, yet much lower values for Norway. Since the $R^2$ increases and probability value decreases for both markets, the statistics imply more evidence for long-run predictability for dividend-growth for the Norwegian market than the U.S. market. In fact, dividend-growth seem to be most forecastable in Norway. Corresponding findings are found in the papers of (2010) and (2018), who found that dividend-growth gives *stronger* statistical evidence in the long-run for European markets than U.S. markets. Return gives *better* forecasts in the longer-horizon for U.S. market compared to the Norwegian market. Moreover, the method of Cochrane (2008) cannot help predicting the Norwegian market for return. Therefore, we cannot see that the dividend-yield model appears to be applicable enough for the Norwegian market for return - something that was also brought up in Engsted and Pedersen (2010) and Monteiro (2018) when analysing the European market on the same premises.

### 5.7 Biases in our findings

Our final table serves only to showcasing the eventual bias(es) in the regression for each coefficient, which appears as the difference in mean of the simulated coefficients and the null-hypotheses Cochrane (2008).

Table 8 shows the means of the simulated coefficients. The possible biases emerge as a result of obtaining a mean lower than the expected value for the corresponding null. For the regressions, the mean of $b_r$ is lower than the null on both values of $\phi$ for both markets, indicating an upwards bias for $b_r$. Dividend growth, on the other hand, is considered biased downwards due to the expected value from the null being higher in relation to the mean for both markets. This was expected due to the fact that dividend growth autocorrelation $\phi$ is a near-unit-root process. Interestingly enough, a larger $\phi$ tends to incur an increasing bias as a result of a higher $\phi$; all of which applies to the long-run

| | | | $\beta_r$ | $\phi$ | $\phi$ | $\beta_r^{lr}$ | $\beta_d^{lr}$ |
|---|---|---|---|---|---|---|---|
| **Norwegian 1984-2018** | 0.658729 | Null | 0.000 | -0.360 | 0.659 | 0.000 | -1.000 |
| | | Mean | 0.049 | -0.397 | 0.571 | 0.057 | -0.943 |
| | 0.990000 | Null | 0.000 | -0.038 | 0.990 | 0.000 | -1.000 |
| | | Mean | 0.078 | -0.098 | 0.848 | 0.416 | -0.584 |
| **U.S 1984-2018** | 0.878752 | Null | 0.000 | -0.141 | 0.879 | 0.000 | -1.000 |
| | | Mean | 0.080 | -0.175 | 0.762 | 0.257 | -0.743 |
| | 0.990000 | Null | 0.000 | -0.033 | 0.990 | 0.000 | -1.000 |
| | | Mean | 0.097 | -0.075 | 0.848 | 0.466 | -0.534 |

Table 9: Mean values of coefficients and long-run

The means in table 7 shows the mean of the simulated corresponding coefficient. The null-hypothesis represents the null for each coefficient.

forecast for $B_r$ and $B_d$. Moreover, The rejections of $br^{lr}$ = 0 or $bd^{lr}$ = -1, we are still able to distinguish the biased null value $br^{lr}$. For the Norwegian market, we are able to distinguish even more for the $b_r^{lr}$, which cannot be rejected under the unforecastable null-hypothesis for return (identity 10).

We calculated how much the biases accounts for the sample value by taking the difference between the simulated mean of the coefficient and the null-hypothesis for the corresponding $\phi$. Therefore, the small bias of 0.057 for $b_r^{lr}$ and the small bias of $b_d^{lr} \approx 0.06$ accounts for almost nothing of their sample values, which supports the evidence that the parameter estimates will be close to their sample values $b_r^{lr}$= 0.34 $b_d^{lr} = -0.66$ for Norway. Due to a substantially higher difference between the simulated mean and the value for the null-hypothesis for the U.S, and a very low bias for the Norwegian market, the simulated sample mean will therefore be closer to the null-hypothesis for the Norwegian market than the U.S. This means that the Norwegian market gives stronger evidence to *not reject* the unforecastable null-hypothesis for return and less evidence of *rejecting* the unforecastable return null-hypothesis for

U.S. market. Since this supports the long-run conclusion from table 6, 7 and 8, we are more convinced that most of the variation in dividend-yield comes from dividend-growth variation in the Norwegian market, while most comes from return variation in the U.S. market. This is consistent with previous literature from Engsted and Pedersen (2010) and Monteiro (2018), which is that the dividend-yield variation comes most from return in the U.S. market and most from dividend-growth in the Norwegian market.

### 5.8 Extension of Cochrane's data

Regarding the replication of Cochrane's data, we found very little worth mentioning in terms of difference and have therefore chosen to put all Tables in the appendix Extension. All Tables with the sample period 1926-2004 are replications of Cochrane (2008) and present exactly the same results as those in his article. However, the Tables containing probability values for 1926-2004 are marginally different from those in his article due to random draws in the calculations. On top of this, we then extended the original dataset by adding 14 extra years.

Initially, Cochrane started by regressing the one-period return and dividend growth using 5 different regressions in the second Table. His strict hypotheses were based on specifying a null-hypothesis that assumes that return is unforecastable and another assuming that the dividend-growth is unforecastable, respectively. He show that the return is significant, while dividend growth is not significant. However, in regards to the Monte Carlo in Table 13 (Cochrane's Table 3), he states the following: *"(...) the lack of dividend forecastability in the data gives far stronger statistical evidence against the null than does the presence of return"* Cochrane (2008, p. 1543). Table 13 shows the Monte Carlo Simulation with probabilities, interpreted with bootstrapping-inference. From what we can see, the probability of obtaining a higher simulated value than the mean is a lot higher for return coefficient $B_r$ as opposed to the dividend-growth coefficient $B_d$ for the U.S. The original dataset and the extended dataset reveal exactly the same interpretation in Table 10, except for the different sign of the dividend growth coefficient. Moreover, the change in sign means that the interpretation actually aligns with the *economic intuition*. Cochrane expected a negative dividend coefficient, rather than the positive coefficient that he found (Cochrane (2008, p. 1535)). So far it is unknown why the sign is different between these two datasets for the U.S. The $\phi$ is also the same for the original and extended dataset, which is one of the reasons of why the results, regardless of the different starts and endings of the samples. Despite the financial crisis in 2008 (which was not included in Cochrane's article), the statistics show almost no difference except skewness and kurtosis. The distribution remained almost the same for all variables, except for dd and dp, which generated a higher t-stat for the joint test and thereby moved further away from a normal distribution.

Moving on to Table 12: increasing phi, an increase in phi makes almost no difference between the original and the extended dataset except for excess return. The probability values stay almost the same for $b_r$ and $b_d$ for U.S. 1926-2004 and U.S. 1926-2018. Moreover, the long-run estimates in Table 14: Long-run horizon, are almost the same, where $b_r^{lr} \approx 1$ and $b_d^{lr} \approx 0$ for both datasets, with almost the same t-statistics and % probability values. The long-run forecasting power in Table 15 shows that the probability values for return decrease as k increases. However, the extended dataset generates a higher probability value for return than the original dataset. Does this mean that using a longer sample that includes a financial crisis and more volatility, leads to even less reliable long-run forecasting? It is tempting to claim that the longer sample gives a *better* estimate of whether or not long-run forecasting is reliable. Finally, as expected, the implied probabilities stay almost the same between the two datasets.

In sum, we obtained the same interpretation as Cochrane for both his original datasets an our extended datasets. He concluded that a forecastable null-hypothesis for dividend growth gives more evidence against the null, than the unforecastable return (Cochrane (2008)) for the one-period regression. Meaning for the one-period, return is predicted by dividend yield and there is no predictability in dividend growth. The same interpretation can be seen in the extended sample as well, where the probability for $B_d$ in the Monte Carlo Simulation is much lower than the return. However, when using the long-run regression, the extended sample provides a different interpretation as the original dataset. We find less predictability over the long-run than the original dataset. The probability values are therefore higher, and $R^2$ is lower for the original- and extended dataset.

## 6 Conclusion

For the time-varying stock market return, during the period 1984-2018, the Norwegian market gives *stronger* evidence against the unforecastable null-hypothesis for dividend growth than return, while the U.S market finds the opposite conclusion, during the period 1984-2018. The estimates obtained in table 6 imply that most of the variation in dividend yield comes more from the variation of dividend-growth and less from the variation of return in the Norwegian market. That is, the variation in dividend-yield is accounted for in a higher degree for return compared to the variation in dividend-growth.

The $R^2$ in table 7 shows that the explanatory power for return in the long-run increases for the U.S. market, while it decreases for the Norwegian market. With respect to the U.S. market, this is equivalent to the findings of Fama and French (1988), who also obtained an increasing $R^2$ for long-run return for the U.S. The long-run dividend-growth, however, shows the opposite pattern. That is, the $R^2$ for dividend-growth for the Norwegian data increases as we increase the number of lags, and the $R^2$ decreases for the U.S. This is is exactly what was found by Engsted and Pedersen (2010), who found that $R^2$ increased for dividend growth in the long-run for the *European* market and decreases in the U.S.

The results imply that Cochrane's VAR system (2008) can only yield increasing long-run coefficients for return in the U.S., not for Norwegian Stock Market. The autocorrelation for dividend-yield, $\phi_{Norway} = 0.659$ is quite different from $\phi_{U.S.} = 0.879$. Fama and French (1988) concluded that "high autocorrelation causes the variance of expected returns to grow faster than the return horizon" (Fama and French (1988, p. 1)). Due to the low $\phi_{Norway}$, which gives a low and negative return coefficient, while the larger $\phi_{U.S}$ causes a higher return for the

U.S. market, we can say that the findings are consistent with the results from Fama and French (1988) for each market.

The simulations give higher probability values than the sample mean for the return in the Norwegian data compared to the U.S. data. We find that dividend-growth gives *stronger* long-run predictability than return for the Norwegian market. We also find that return is *more* forecastable in the long-run than dividend growth in the U.S. market. As we can see, this result is different for the two markets. This is consistent with what Engsted and Pedersen (2010) concluded with, who finds better long-run predictability for return in U.S. than the European market and the opposite for the dividend growth. The same applies for Monteiro (2018), who also arrived at different conclusions between the U.S. and European markets; that is, return provides stronger evidence for forecastability in the U.S., but the opposite for Germany, France, and Italy. The replicated results of Cochrane's Sample (1926-2004), the additional 14 years of extended sample data *and* the U.S. data 1984-2018 gives approximately the same $\phi$ and results. When comparing the U.S market and Norwegian market for 1984-2018 on the other hand, the results are quite different. Hence, predictability in dividend-growth and return estimates with the dividend-yield model from Campbell and Shiller (1988) *depends* on which market we forecast, by using the method from Cochrane (2008).Hence, using the Cochrane System (2008) as applied to Campbell and Shiller's dividend-yield model (1988) predictability in dividend-growth and return estimates are very dependent on which markets we forecast, and shows more accuracy when used for long-run, rather than short-term estimates.

Future research could examine the VAR-system with more independent variables instead of only using one. This could be developed to a factor model, by using book-to-market ratios, size and so on. Alternatively, assuming that the indices being used has more data, one could use this additional data in order to forecast for, at most, 11 years as we did. An out-of-sample test with many years of observations is also an interesting avenue for future research.

**APPENDIX**

# A Cochrane replication and extension- tables

In this appendix section, the tables for U.S 1926-2004, which is fully replicated from Cochrane's article and U.S 1926-2018, is presented here. Only the tables are presented, whereas comparison between each datasets are done in results and analysis.

| | | Correlation, std on diagonal | | | | | Other statistics | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r** | **dd** | **dp** | **RF** | **CPI** | **skew** | **kurt** | **jointly** | **mean** |
| **US 1926-2004** | **r** | 0.20 | 0.65 | 0.20 | -0.07 | -0.17 | -0.68 | 0.28 | 30.52 | 0.07 |
| | **dd** | 0.65 | 0.14 | 0.02 | -0.10 | -0.12 | -0.02 | 0.17 | 26.29 | 0.01 |
| | **dp** | 0.20 | 0.02 | 0.41 | -0.18 | -0.01 | -0.94 | 0.93 | 25.65 | -3.27 |
| | **RF** | -0.07 | -0.10 | -0.18 | 0.04 | 0.39 | 1.01 | 1.16 | 24.62 | 0.04 |
| | **CPI** | -0.17 | -0.12 | -0.01 | 0.39 | 0.04 | 0.21 | 3.03 | 0.60 | 0.03 |
| **U.S 1926-2018** | **r** | 0.20 | 0.63 | 0.20 | -0.03 | -0.15 | -0.80 | 0.68 | 30.80 | 0.06 |
| | **dd** | 0.63 | 0.14 | -0.01 | -0.09 | -0.16 | 0.06 | -0.02 | 35.31 | 0.02 |
| | **dp** | 0.20 | -0.01 | 0.44 | 0.06 | 0.07 | -0.45 | -0.33 | 46.12 | -3.33 |
| | **RF** | -0.03 | -0.09 | 0.06 | 0.03 | 0.46 | 1.13 | 1.68 | 26.46 | 0.03 |
| | **CPI** | -0.15 | -0.16 | 0.07 | 0.46 | 0.04 | 0.35 | 3.84 | 4.56 | 0.03 |

Table 10: Statistics (Extension)

r is log return t+1 and dd is log dividend change t+1, deflated by CPI. Dp is log dividend price ratio at time t, RF is T-bill 3 month at t+1 and CPI is consumption price index at t+1. In "Correlation, std on diagonal", the diagonal is standard deviation of the corresponding letters and the rest are correlation

| | Regression | $\beta$ | se | $t(\beta)$ | $R^2$ | Stdx*b |
|---|---|---|---|---|---|---|
| **US 1926-2004** | $R_{t+1} = \alpha_R + \beta_R(D_t/P_t) + \varepsilon^R_{t+1}$ | 3.387 | 1.488 | 2.28 | 0.0582 | 0.049 |
| | $R_{t+1} - R^f_t = \alpha_R + \beta_{R-\text{RF}}(D_t/P_t) + \varepsilon^R_{t+1}$ | 3.829 | 1.465 | 2.61 | 0.0737 | 0.056 |
| | $D_{t+1}/D_t = \alpha_D + \beta_D(D_t/P_t) + \varepsilon^D_{t+1}$ | 0.073 | 1.159 | 0.06 | 0.0001 | 0.001 |
| | $r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$ | 0.097 | 0.050 | 1.92 | 0.0398 | 0.040 |
| | $\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1}$ | 0.008 | 0.044 | 0.18 | 0.0005 | 0.003 |
| **U.S 1926-2018** | $R_{t+1} = \alpha_R + \beta_R(D_t/P_t) + \varepsilon^R_{t+1}$ | 2.886 | 1.169 | 2.47 | 0.0532 | 0.046 |
| | $R_{t+1} - R^f_t = \alpha_R + \beta_{R-\text{RF}}(D_t/P_t) + \varepsilon^R_{t+1}$ | 2.980 | 1.159 | 2.57 | 0.0573 | 0.047 |
| | $D_{t+1}/D_t = \alpha_D + \beta_D(D_t/P_t) + \varepsilon^D_{t+1}$ | -0.164 | 0.998 | -0.16 | 0.0003 | 0.003 |
| | $r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$ | 0.087 | 0.042 | 2.07 | 0.0382 | 0.038 |
| | $\Delta d_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1}$ | -0.002 | 0.035 | -0.06 | 0.0000 | 0.001 |

Table 11: Forecasting (Extension)

Table 11 shows the probability values that the simulated coefficients are larger than the corresponding coefficients for different values of $\phi$. The probability value 25.63 for $b_r$ is calculated as $\frac{\sum b_r^{sim} > b_r^{data}}{50.000}$ using $\phi = 0.659$ in the simulations. The coefficient for return is estimated under the unforecastale return-null *while* dividend growth is. The coefficient for dividend-growth is estimated under the null that assumes that dividend-growth cannot be forecasted *while* return is forecastable.

| | | Estimates | | | ε s. d. diagonal and corr | | | Null 1 | Null 2 |
|---|---|---|---|---|---|---|---|---|---|
| | | $\hat{\beta}$ | $\sigma(\hat{\beta})$ | **Implied** | **r** | **Δd** | **dp** | $\beta, \Phi$ | $\beta, \Phi$ |
| **US 1926-2004** | **r** | 0.097 | 0.050 | 0.101 | 0.196 | 0.660 | -0.700 | 0.000 | 0.000 |
| | **Δd** | 0.008 | 0.044 | 0.004 | 0.660 | 0.140 | 0.075 | -0.093 | -0.046 |
| | **dp** | 0.941 | 0.047 | 0.945 | -0.700 | 0.075 | 0.153 | 0.941 | 0.990 |
| **U.S 1926-2018** | **r** | 0.087 | 0.042 | 0.088 | 0.192 | 0.640 | -0.679 | 0.000 | 0.000 |
| | **Δd** | -0.002 | 0.035 | -0.003 | 0.640 | 0.142 | 0.130 | -0.091 | -0.044 |
| | **dp** | 0.942 | 0.039 | 0.943 | -0.679 | 0.130 | 0.154 | 0.942 | 0.990 |

Table 12: VAR (Extension)

Each row represents the one-period regression from the VAR system described in methodology. The first row, for instance, use the regression $r_{t+1} = \alpha + \beta_r(d_t - p_t) + \varepsilon^r_{t+1}$, equivalent to identity 1; the same applies for the second and third row for each set. The implied values are calculated by solving for the corresponding dependent variable in identity 5. $\rho$ is defined as the constant taylor-approximated point estimate for the dividend yield and is used to calculate the implied value for the corresponding dependent variable in in table 3 above. The null columns are the coefficients which are used in the simulations under the null.

| | | Percent probability values | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Real returns | | | | Excess returns | | | |
| | $\phi$ | $\beta_r$ | $\phi$ | $br^{lr}_{min}$ | $br^{lr}_{max}$ | $\beta_r$ | $\phi$ | $br^{lr}_{min}$ | $br^{lr}_{max}$ |
| **US 1926-2004** | **0.900** | 23.50 | 0.67 | 0.36 | 0.62 | 18.87 | 0.42 | 0.10 | 0.17 |
| | **0.941** | 22.73 | 1.81 | 1.30 | 1.88 | 17.70 | 1.18 | 0.44 | 0.70 |
| | **0.960** | 22.86 | 3.14 | 2.50 | 3.32 | 16.66 | 1.70 | 0.85 | 1.24 |
| | **0.980** | 21.42 | 4.74 | 4.12 | 5.31 | 16.26 | 2.72 | 1.85 | 2.42 |
| | **0.990** | 21.62 | 6.42 | 5.78 | 7.37 | 16.11 | 3.64 | 2.62 | 3.55 |
| | **1.000** | 21.78 | 8.98 | 8.54 | 10.40 | 16.78 | 4.64 | 3.70 | 5.04 |
| | **1.010** | 18.36 | 10.25 | 10.50 | 12.58 | 14.20 | 5.33 | 5.32 | 6.58 |
| | **Draw** | 25.02 | 1.80 | 1.56 | 2.03 | 20.68 | 1.10 | 0.80 | 1.10 |
| **U.S 1926-2018** | **0.900** | 21.25 | 0.46 | 0.51 | 0.58 | 20.36 | 0.46 | 0.48 | 0.55 |
| | **0.942** | 19.11 | 1.44 | 1.57 | 1.74 | 18.77 | 1.33 | 1.37 | 1.53 |
| | **0.960** | 18.82 | 2.66 | 2.84 | 3.08 | 17.43 | 2.42 | 2.47 | 2.75 |
| | **0.980** | 18.32 | 4.90 | 5.04 | 5.42 | 17.15 | 4.49 | 4.54 | 4.88 |
| | **0.990** | 18.36 | 6.75 | 6.86 | 7.30 | 17.34 | 6.14 | 6.23 | 6.66 |
| | **1.000** | 17.98 | 9.73 | 9.73 | 10.28 | 16.80 | 8.47 | 8.46 | 9.06 |
| | **1.010** | 14.14 | 11.41 | 11.16 | 11.81 | 13.26 | 10.46 | 10.33 | 10.86 |
| | **Draw** | 22.78 | 1.91 | 1.93 | 2.13 | 22.08 | 1.82 | 1.85 | 2.02 |

Table 13: Increasing $\phi$ (phi) (Extension)

Table 13 shows the probability values that the simulated coefficients are larger than the corresponding coefficients for different values of $\phi$. The probability value 25.63 for $b_r$ is calculated as $\frac{\sum b_r^{sim} > b_r^{data}}{50.000}$ using $\phi = 0.659$ in the simulations. The coefficient for return is estimated under the unforecastale return-null *while* dividend growth is. The coefficient for dividend-growth is estimated under the null that assumes that dividend-growth cannot be forecasted *while* return is forecastable.

| | | $\beta_r$ | $t_r$ | $\phi$ | $t_d$ |
|---|---|---|---|---|---|
| **US 1926-2004** | **Real** | 22.095 | 10.120 | 1.800 | 1.675 |
| | **Excess** | 17.045 | 6.015 | 1.085 | 0.795 |
| **U.S 1926-2018** | **Real** | 19.455 | 7.140 | 1.620 | 1.685 |
| | **Excess** | 18.570 | 6.515 | 1.505 | 1.520 |

Table 14: Monte Carlo Simulation by Bootstrapping (Extension)

The columns gives the probability that the coefficients are larger than the corresponding coefficients in the sample data. Example is $b_r^{real} = \frac{\sum b_r^{sim(real)} > b_r^{data(real)}}{50.000}$. 50.000 samples of $b_r$, $t_r$, $b_d$ and $t_d$ in a Monte Carlo simulation and calculated how many of these samples that were greater than the sample coefficients. The coefficient for return is estimated under the unforecastale return-null *while* dividend growth is. The coefficient for dividend-growth is estimated under the null that assumes that dividend-growth cannot be forecasted *while* return is forecastable.

| | | $\beta^{lr}$ | se | t | % prob. values |
|---|---|---|---|---|---|
| **US 1926-2004** | $r_{t+1}$ | 1.09 | 0.44 | 2.48 | 1.31-1.83 |
| | $\Delta d_{t+1}$ | 0.09 | 0.44 | 2.48 | 1.31-1.83 |
| | **Excess** $r_{t+1}$ | 1.23 | 0.47 | 2.62 | 0.41-0.68 |
| **U.S 1926-2018** | $r_{t+1}$ | 0.97 | 0.38 | 2.59 | 1.76-1.95 |
| | $\Delta d_{t+1}$ | -0.03 | 0.38 | 2.59 | 1.76-1.95 |
| | **Excess** $r_{t+1}$ | 0.99 | 0.38 | 2.59 | 1.53-1.68 |

Table 15: Long-run (Extension)

Table 15 shows the long-run coefficients. Note that the "lr" superscript denotes long-horizon coefficients. $\hat{\beta}_r^{lr}$ is calculated as $\hat{\beta}_r^{lr} = \frac{\hat{b_r}}{1-\rho\hat{\phi}}$ where $\hat{b_r}$ is the coefficient from the regression $r_{t+1} = \alpha + \hat{\beta_r}(d_t - p_t) + \varepsilon_{t+1}^r$. $\hat{\phi}$ comes from the autocorrelation $d_{t+1} - p_{t+1} = \alpha + \hat{\phi}_{dp}(d_t - p_t) + \varepsilon_{t+1}^{dp}$. The % probability values are the probability range under the null hypotheses $\phi$ from each of the markets. The t statistic for $b_d^{lr}$ is the t stats for the hypothesis that $\hat{b}_d^{lr} = -1$.

| | | Weighted $\sum_{j=1}^{k} \rho^{j-1} r_{t+j} = \alpha + \beta_r^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | | Unweighted $\sum_{j=1}^{k} r_{t+j} = \alpha + \beta_r^{(k)}(d_t - p_t) + \varepsilon_{t+k}$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Direct | | | | Implied | | | Direct | | | | Implied | | |
| | k | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ | $R^2$ | $\hat{\beta}_r^{(k)}$ | $\phi_{\text{data}}$ | $\phi_{99}$ |
| **US 1926-2004** | 1.0 | 0.10 | 22.71 | 21.50 | 0.04 | 0.10 | 22.71 | 21.50 | 0.10 | 22.71 | 21.50 | 0.04 | 0.10 | 22.71 | 21.50 |
| | 5.0 | 0.35 | 28.34 | 28.41 | 0.12 | 0.40 | 16.51 | 18.22 | 0.37 | 28.68 | 29.04 | 0.12 | 0.43 | 16.22 | 18.07 |
| | 10.0 | 0.80 | 15.82 | 15.70 | 0.27 | 0.65 | 10.39 | 14.32 | 0.92 | 15.84 | 15.69 | 0.27 | 0.75 | 9.21 | 13.60 |
| | 15.0 | 1.38 | 4.56 | 4.53 | 0.51 | 0.80 | 6.42 | 11.51 | 1.68 | 5.00 | 4.78 | 0.49 | 0.98 | 4.49 | 10.06 |
| | 20.0 | 1.49 | 4.81 | 4.96 | 0.61 | 0.89 | 4.28 | 9.67 | 1.78 | 7.94 | 7.78 | 0.49 | 1.15 | 2.22 | 7.52 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | 1.04 | 1.79 | 7.34 | ∞ | ∞ | ∞ | ∞ | 1.64 | 0.50 | 9.12 |
| **U.S 1926-2018** | 1.0 | 0.09 | 19.72 | 17.43 | 0.04 | 0.09 | 19.72 | 17.43 | 0.09 | 19.72 | 17.43 | 0.04 | 0.09 | 19.72 | 17.43 |
| | 5.0 | 0.28 | 29.62 | 28.78 | 0.12 | 0.36 | 14.92 | 15.22 | 0.30 | 30.03 | 29.33 | 0.11 | 0.39 | 14.75 | 15.10 |
| | 10.0 | 0.56 | 23.97 | 23.92 | 0.28 | 0.59 | 9.56 | 12.50 | 0.63 | 24.46 | 24.32 | 0.27 | 0.67 | 8.62 | 12.02 |
| | 15.0 | 0.74 | 22.42 | 22.83 | 0.35 | 0.73 | 5.94 | 10.37 | 0.89 | 23.28 | 23.72 | 0.33 | 0.89 | 4.66 | 9.18 |
| | 20.0 | 0.98 | 15.62 | 16.56 | 0.39 | 0.82 | 4.26 | 8.80 | 1.20 | 18.24 | 19.25 | 0.34 | 1.04 | 2.58 | 7.47 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | 0.96 | 1.94 | 7.10 | ∞ | ∞ | ∞ | ∞ | 1.49 | 0.56 | 10.26 |

Table 16: Long-run forecasting (Extension)

Table 16 shows the long-run coefficients for the regression $b_r^{(k)} = \sum_{j=1}^{k} \rho^{j-1}\phi^{j-1} b_r$, where k is the number of lags. The table includes direct and indirect estimates for weighted and unweighted regression, respectively (calculation available in direct and indirect estimates for weighted and unweighted regression). The $\hat{b}_r^{(k)}$ is the coefficient of the long-run estimate for the corresponding lag. The *phi_data* is the probability that the simulated long-run coefficient $\hat{b}_r^{(k)}$ exceeds the sample value $\hat{b}_r^{(k)}$, for each data set.

| | | | $\beta_r$ | $\phi$ | $\phi$ | $\beta_r^{lr}$ | $\beta_d^{lr}$ |
|---|---|---|---|---|---|---|---|
| **US 1926-2004** | 0.940974 | Null | 0.000 | -0.093 | 0.941 | 0.000 | -1.000 |
| | | Mean | 0.049 | -0.097 | 0.886 | 0.246 | -0.754 |
| | 0.990000 | Null | 0.000 | -0.046 | 0.990 | 0.000 | -1.000 |
| | | Mean | 0.057 | -0.050 | 0.927 | 0.398 | -0.602 |
| **U.S 1926-2018** | 0.941731 | Null | 0.000 | -0.091 | 0.942 | 0.000 | -1.000 |
| | | Mean | 0.040 | -0.096 | 0.895 | 0.206 | -0.794 |
| | 0.990000 | Null | 0.000 | -0.044 | 0.990 | 0.000 | -1.000 |
| | | Mean | 0.045 | -0.050 | 0.936 | 0.374 | -0.626 |

Table 17: Biases (Extension)

The means in table 17 shows the mean of the simulated corresponding coefficient. The null represents the null for each coefficient.

# B Explanations

## B.1 Stationarity in time series

The stationarity of a time series is a characteristic that can strongly influence its behaviour and properties. For the sake of illustration, the word ‘shock’ is commonly used to express a change or an unexpected change in a variable or perhaps simply the value of the error term during a particular time period. As explained in Brooks (2014):

”For a stationary series, ‘shocks’ to the system will gradually die away. That is, a shock during time t will have a smaller effect in time t +1, a smaller effect still in time t +2, and so on. This can be contrasted with the case of non-stationary data, where the persistence of shocks will always be infinite, so that for a non-stationary series, the effect of a shock during time t will not have a smaller effect in time t +1, and in time t +2, etc.”. Brooks (2014, p. 319).

## B.2 Blue assumptions

What is referred to as BLUE assumptions are prerequisites for the OLS estimator that we expect to to hold, to ensure that the parameter estimates has the desirable underlying properties for accurate statistical use. ”If assumptions 1–4 hold, then the estimators $\alpha$ and $\hat{\beta}$ determined by OLS will have a number of desirable properties, and are known as best linear unbiased estimators (BLUE)” (Brooks (2014, p. 44)). Here are definitions of the blue assumptions, described in Brooks (2014):

- ”‘Best’ – means that the OLS estimator $\hat{\beta}$ has minimum variance among the class of linear unbiased estimators; the Gauss–Markov theorem proves that the OLS estimator is best by examining an arbitrary alternative linear unbiased estimator and showing in all cases that it must have a variance no smaller than the OLS estimator." (Brooks (2014, p. 45))

- "'Estimator' – $\hat{\alpha}$ and $\hat{\beta}$ are estimators of the true value of $\alpha$ and $\beta$" (Brooks (2014, p. 45))
- "Linear' – $\hat{\alpha}$ and $\hat{\beta}$ are linear estimators – that means that the formula for $\hat{\alpha}$ and $\hat{\beta}$ are linear combinations of the random variables." (Brooks (2014, p. 45))
- "'Unbiased' – on average, the actual values of $\hat{\alpha}$ and $\hat{\beta}$ will be equal to their true values" (Brooks (2014, p. 45))

The assumptions for a the linear estimators are the following:

1. E($u_t$)=0 - The expectation of residuals observed today is, on average, equal to zero.
2. var($u_t$) = $\sigma^2 > \infty$ - Assuming that the variance of residuals are constant, which implies homoscedasticity.
3. cov($u_i, u_j$) $\neq 0$ for i $\neq$ j - No covariance between error terms over time. Also know as being serially uncorrelated or linearly independent of one another.
4. cov($u_i, x_t$) $\neq 0$ - There is no relationship between the error term and the x variable
5. ($u_t$)  N (0,$\sigma^2$) - The error term is normally distributed. In principal, this is a joint assumption of assumption 1 and 2.

### B.3 Boostrapping

Bootstrapping is related to simulation, but with one crucial difference:

"With simulation, the data are constructed completely artificially. Bootstrapping, on the other hand, is used to obtain a description of the properties of empirical estimators by using the sample data points themselves, and it involves sampling repeatedly with replacement from the actual data" (Brooks (2014, p. 553))

According to Brooks (2014), the advantage of bootstrapping over the use of analytical results is that it allows the researcher to make inferences without making strong distributional assumptions, since the distribution employed will be that of the actual data. Instead of imposing a shape on the sampling distribution of the $\hat{\phi}$ value, bootstrapping involves empirically estimating the sampling distribution by looking at the variation of the statistic within-sample.

### B.4 Direct and indirect

In table 6, we are using weighted and unweighted regression, where the weighted use $\rho$ and the unweighted does not include $\rho$.

### B.5 Monte Carlo Simulation

The Monte Carlo Simulations are calculated with random draws using the bootstrap method (Brooks (2014)).

# C Assumptions of linear regression

In this section, we present the results from the assumptions on linear regressions. We briefly explain the assumptions, if the assumptions are violated, and how it might affect the U.S and Norwegian Market data, respectively. Due to this section being relevant to only some aspects of the paper and a bit long, we have only decided only to include the test results for the BLUE assumptions below. We have presented a summary table of the variables that have either satisfied or violated the assumptions before taking a closer look at each assumption. Conclusively, we have written an overall summary for each market. We have followed the steps of the test in Python, which was described by Macaluso (2018) and the theory provided in the book of Brooks (2014).

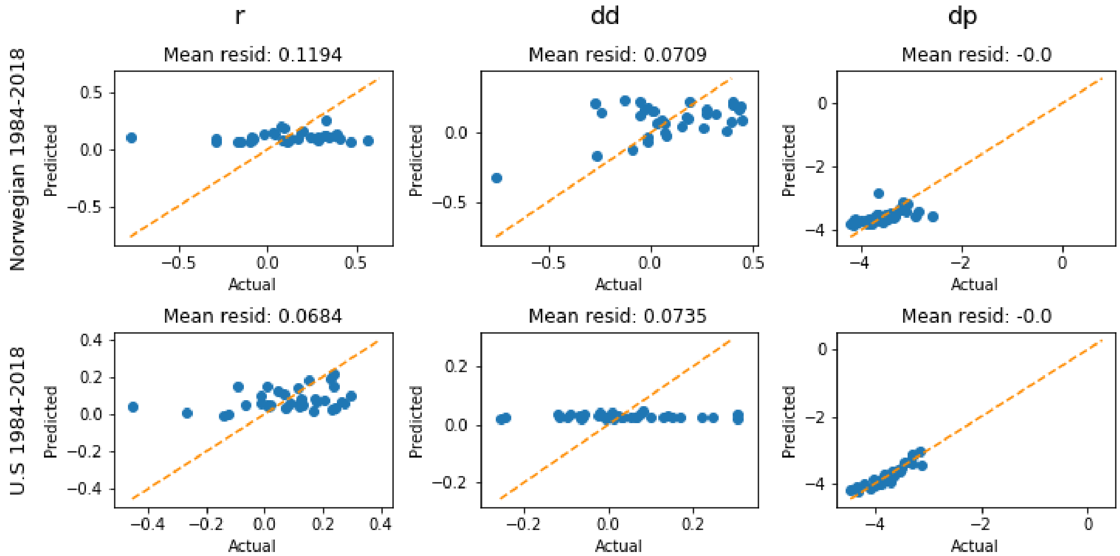
## C.1 BLUE assumptions

| | | **Assumption 1** | **Assumption 2** | **Assumption 3** | **Assumption 4** | **Assumption 5** |
|---|---|---|---|---|---|---|
| **Norwegian 1984-2018** | **r** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| | **dd** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| | **dp** | Satisfied | Satisfied | Satisfied | Satisfied | Satisfied |
| **U.S 1984-2018** | **r** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| | **dd** | Satisfied | Not satisfied | Positive | Satisfied | Not satisfied |
| | **dp** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |

Table 17: Summary of BLUE assumptions

The summary of assumption is a table which summarizes the variables that either satisfies or violates each of the BLUE assumptions. The first panel represents the Norwegian market, and the second panel represents the U.S market.

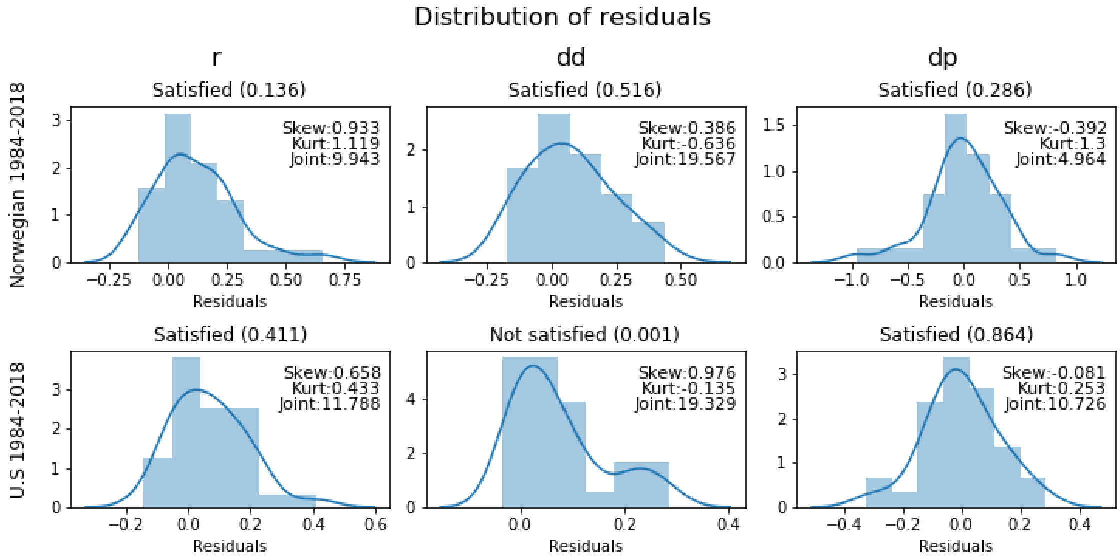
## C.2 Assumption 1




Assumption 1: Linearity

Assumption 1 tests if the mean of the data is equal to zero. According to the data, the mean of the residuals is insignificantly different from zero for the Norwegian and U.S market. Hence, assumption 1 holds both all identities.

## C.3 Assumption 2




Assumption 2: Normality

The summary of assumption is a table which summarizes the variables that either satisfies or violates each of the BLUE assumptions. The first panel represents the Norwegian market, and the second panel represents the U.S market.

**U.S** Assumption 2 tests if the mean of the residuals are equal to zero. According to the data, assumption 2 is holds for all identities, but identity 2 for the U.S market. Identity 2 is the dependent dividend growth identity. This indicates heteroscedasticity and sign of inconstant variance for identity 2. The results are derived from an Anderson-Darling test.

**Norwegian** Assumption 2 tests if the mean of the residuals are equal to zero. According to the data, assumption 2 holds for all identities. This indicates homoscedastic error terms. The results are derived from an Anderson-Darling test.

### C.4 Assumption 3 & 4

| | | Assumption 3 | | Assumption 4 | |
|---|---|---|---|---|---|
| | | Durbin test | Autocorr. | Covariance | Sign of cov? |
| **Norwegian 1984-2018** | **r** | 1.0777 | Positive | -0.001 | No sign |
| | **dd** | 1.3878 | Positive | 0.001 | No sign |
| | **dp** | 2.0121 | No sign | -0.000 | No sign |
| **U.S 1984-2018** | **r** | 1.0410 | Positive | -0.000 | No sign |
| | **dd** | 1.2489 | Positive | -0.000 | No sign |
| | **dp** | 1.4255 | Positive | -0.000 | No sign |

Table 18: Assumption 3 and 4

The table above shows the t-test and the results from a Durbin-Watson test (assumption 3), as well as the covariance between the respective variable and the error term (assumption 4

**U.S 1984-2020**

Assumption 3 tests for autocorrelation in the error-term. For the U.S data, assumption three is violated for all identities . The results indicate signs of positive autocorrelation in the error terms. The results are derived from a Durbin-Watson test. Assumption 4 tests for covariance between the corresponding x variate and the error term. According to the data, the error terms shows signs of no correlation between the error term and the independent variable on any of the identities. Thus, assumption 4 holds all identities. The results are derived directly from the python script by calculating the covariance.

**Norwegian 1984-2020**

Assumption three tests for autocorrelation in the error-term. Assumption three only holds for identity 3, while the other identities shows sign of positive autocorrelation. The results are derived from a Durbin-Watson test. Assumption 4 tests for covariance between the corresponding x variate and the error term. According to the data, the error terms shows signs of no correlation between the error term and the independent variable on any of the identities. Thus, assumption 4 holds all identities. The results are derived directly from the python script by calculating the covariance.

### C.5 Assumption 5

| | | Jarque Bera | Satisfied? |
|---|---|---|---|
| **Norwegian 1984-2018** | **r** | 9.94 | Not satisfied |
| | **dd** | 19.57 | Not satisfied |
| | **dp** | 4.96 | Satisfied |
| **U.S 1984-2018** | **r** | 11.79 | Not satisfied |
| | **dd** | 19.33 | Not satisfied |
| | **dp** | 10.73 | Not satisfied |

Table 19: Assumption 5

The summary of assumption is a table which summarizes the variables that either satisfies or violates each of the BLUE assumptions. The first panel represents the Norwegian market, and the second panel represents the U.S market.

Assumption 5 assumes joint normality. The results are derived from a Barque-Jera test, that tests for excess skewness- and kurtosis in the distribution Brooks (2014). In principal, the tests works as a joint test of assumption 1 and 2 together. . In principal, the tests work

Assumption 5 assumes joint normality. The results are derived from a Barque-Jera test, that tests for excess skewness- and kurtosis in the distribution Brooks (2014). In principal, the tests works as a joint test of assumption 1 and 2 together. Thus, since these assumptions holds for our data, it is not surprising that assumption 5 holds.

### C.6 Summary

U.S 1984-2018

From what we can see, assumption 1 and 4 are the only assumptions that holds for all identities for the U.S. The Anderson-test, that tests for homoscedasticity, holds for all identities but identity 2, which implies heteroscedasticity in the error term for dividend-growth. The Durbin-Watson test shows signs of positive autocorrelation with residuals for all identities, which has makes the standard errors biased downwards relative to the true value Brooks (2014). Lower standard errors gives a higher t-stat than the true value, which can result in conducting a type-I Brooks (2014). $R^2$ is also likely to inflated due to this violation, according to Brooks (2014). Apart from that, the mean in assumption 1 is insignificantly different from zero, the independent variables are uncorrelated with the error term. If assumption 4 was violated, the estimates would not even be consistent as the sample size increases.

Norwegian 1984-2018

From what we can see, assumption 1-2 and 4 holds for all identities, as well as assumption 5 for the DP-ratio. The violation of assumption 3 implies a downwards bias in standard errors relative to the true value Brooks (2014). Lower standard errors gives a higher t-stat than the true value, which can result in conducting a type-I error when doing hypothesis testing. $R^2$ is also likely to inflated due to this violation, according to Brooks (2014).

Note that only assumption 1-4 must hold in order to say that the parameters are, statistically, BLUE Brooks (2014).

| | | Assumption 1 | Assumption 2 | Assumption 3 | Assumption 4 | Assumption 5 |
|---|---|---|---|---|---|---|
| **US 1926-2004** | **r** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| | **dd** | Satisfied | Not satisfied | Positive | Satisfied | Not satisfied |
| | **dp** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| **U.S 1926-2018** | **r** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |
| | **dd** | Satisfied | Not satisfied | Positive | Satisfied | Not satisfied |
| | **dp** | Satisfied | Satisfied | Positive | Satisfied | Not satisfied |

Table 20: Summary of BLUE assumptions (Cochrane)

The summary of assumption is a table which summarizes the variables that either satisfies or violates each of the BLUE assumptions. The first panel represents the U.S. market 1926-2004, and the second panel represents the U.S. market 1926-2018.

# D Models

## D.1 Campbell and Shiller decomposition

Start with the return equation,

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} \tag{1}$$

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} * \frac{\frac{1}{D_t}}{\frac{1}{D_t}} \tag{2}$$

$$R_{t+1} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}})\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}} \tag{3}$$

Loglineralizing equation (3)

$$r_{t+1} = log[1 + e^{(p_{t+1} - dt+1)}] + log(\frac{d_{t+1}}{d_t}) - log(\frac{p_t}{d_t}) \tag{4}$$

$$r_{t+1} = log[1 + e^{(p_{t+1} - dt+1)}] + \Delta d_{t+1} - (p_t - d_t) \tag{5}$$

A function can be approximated using Taylor series (July Thomas and Samir Khan and Jimin Khim (2019)),

$$f(x) \approx f(a) + f'(x) * (x - a) + \frac{1}{2} f''(x) * (x - a)^2 + ... + \frac{1}{n!} f^{(n)}(x) * (x - a)^n \tag{6}$$

Campbell and Shiller (1988) used first order Taylor approximation for the first term in equation (5), $log[1 + e^{(p_{t+1} - d_{t+1})}]$,

$$log[1 + e^{(p_{t+1} - d_{t+1})}] \approx log[1 + e^{p_t - d_t}] + \frac{e^{p_t - d_t}}{1 + e^{p_t - d_t}} * ((p_{t+1} - d_{t+1}) - (p_t - d_t)) \tag{7}$$

Where $x = p_{t+1} - d_{t+1}$ and $a = p_t - d_t$. From here, the approximation is plugged into equation (5),

$$
\begin{aligned}
r_{t+1} &= log[1+e^{p_t-d_t}] + \frac{e^{p_t-d_t}}{1+e^{p_t-d_t}} * ((p_{t+1}-d_{t+1})-(p_t-d_t)) + \Delta d_{t+1} - (p_t-d_t) \\
&\approx log[1+e^{p_t-d_t}] - (p_t-d_t) * \frac{e^{p_t-d_t}}{1+e^{p_t-d_t}} + \frac{e^{p_t-d_t}}{1+e^{p_t-d_t}} * (p_{t+1}-d_{t+1}) + \Delta d_{t+1} - (p_t-d_t) \\
&\approx k + \rho(p_{t+1}-d_{t+1}) + \Delta d_{t+1} - (p_t-d_t)
\end{aligned}
$$

Where $\rho = \frac{e^{p_t-d_t}}{1+e^{p_t-d_t}}$ and k=constant, is dropped, equation (10) is optained.

### D.2 Identities from VAR representation

Start with the VAR system in (Cochrane (2008)):

$$r_{t+1} = \alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1} \tag{1}$$

$$\Delta_{t+1} = \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1} \tag{2}$$

$$d_{t+1} - p_{t+1} = \alpha_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1} \tag{3}$$

Plugging the VAR representation into Campbell and Shiller (1988) linerization:

$$r_{t+1} = \rho(p_{t+1} - d_{t+1}) + \Delta_{t+1} - (p_t - d_t) \tag{4}$$

This gives the equation

$$\alpha_r + \beta_r(d_t - p_t) + \varepsilon^r_{t+1} = -\rho(\alpha_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1}) + \alpha_d + \beta_d(d_t - p_t) + \varepsilon^d_{t+1} - (p_t - d_t)$$

$$\alpha_r + \rho\alpha_{dp} - \alpha_d = -\rho\phi(d_t - p_t) + \beta_d(d_t - p_t) + (d_t - p_t) - \beta_r(d_t - p_t) - \rho\varepsilon^{dp}_{t+1} - \varepsilon^r_{t+1} + \varepsilon^d_{t+1}$$

$$\alpha_r + \rho\alpha_{dp} - \alpha_d = (d_t - p_t) * (-\rho\phi + \beta_d + 1 - \beta_r) + (-\rho\varepsilon^{dp}_{t+1} - \varepsilon^r_{t+1} + \varepsilon^d_{t+1})$$

From the last equation, we get the same identity for regression coefficients and the error link as in Cochrane (2008):

$$\beta_r = 1 - \rho\phi + \beta_d \tag{5}$$

$$\varepsilon^r_{t+1} = \varepsilon^d_{t+1} - \rho\varepsilon^{dp}_{t+1} \tag{6}$$

### D.3 $B^{lr}$

Using the Campbell and Shiller (1988) present value theorem, gives the identity

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta_{t+j} \tag{7}$$

Multiply with $(d_t - p_t) - E(d_t - p_t)$ and take the expectation, gives

$$var(d_t - p_t) = cov\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) - cov\left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta_{t+j}, d_t - p_t \right) \tag{8}$$

Dividing by $var(d_t - p_t)$, the equation becomes

$$\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) - \beta\left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta_{t+j}, d_t - p_t \right) = 1 \tag{9}$$

For $\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right)$ it becomes

$$\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) = \sum_{j=1}^{\infty} \rho^{j-1} \beta\left( r_{t+j}, d_t - p_t \right) \tag{10}$$

$$\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_r \tag{11}$$

$$\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) = \frac{b_r}{1 - \rho\phi} \tag{12}$$

$$\beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) = b_r^{lr} \tag{13}$$

The same can be accomplished using $b_d^{lr}$ and we finally get the equation,

$$b_r^{lr} - b_d^{lr} = 1 \tag{14}$$

## E Inconsistencies

In Cochrane (2008), there are inconsistency in table 6. In his paper, the unweighted implied coefficients are 1.02, 1.26 and 1.41 for 10, 15 and 20 years of forecast, respectively. These coefficients should be 0.75, 0.98 and 1.15 instead, because unweighted $b_r$ is calculated as $b_r^{(k)} = b_r \frac{1-\rho^k \phi^k}{1-\rho\phi}$ in his paper, while they should not contain $\rho$. This mistake is explained by Cochrane (2020). Also, he used the word "power" in terms of probabilities s incorrect in this paper.

## Reference

Bernt. 2020. OBI Financial data at BI. Accessed August 3, 2020. http://finance.bi.no/~bernt/obidata/obidata/?fbclid=IwAR28pBjc2E3jcCy1UVCXdcB1ys5qzFhj77GYKcKyMh76r9DU8uEmhtMqWyA.

Bloomberg. 2020. Bloomberg.com. Accessed August 22, 2020. https://www.bloomberg.com/europe.

Brooks, Chris. 2014. Introductory Econometrics for Finance. doi:10.1017/cbo9781139540872.

Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." Review of Financial Studies. ISSN: 0893-9454. doi:10.1093/rfs/1.3.195.

Campbell, John Y., and Samuel B. Thompson. 2008. "Predicting excess stock returns out of sample: Can anything beat the historical average?" Review of Financial Studies 21, no. 4 (July): 1509–1531. ISSN: 08939454. doi:10.1093/rfs/hhm055. https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhm055.

Cochrane, John H. 2008. "The dog that did not bark: A defense of return predictability." Review of Financial Studies 21 (4): 1533–1575. ISSN: 08939454. doi:10.1093/rfs/hhm046. https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhm046.

———. 2020. John H. Cochrane - Research. Accessed August 2. https://faculty.chicagobooth.edu/john.cochrane/research/.

Engsted, Tom, and Thomas Q. Pedersen. 2010. "The dividend–price ratio does predict dividend growth: International evidence." Journal of Empirical Finance 17 (4): 585–605. ISSN: 09275398.

doi:10.1016/j.jempfin.2010.01.003. https://linkinghub.elsevier.com/retrieve/pii/S0927539810000113.

Fama, Eugene F., and Kenneth R. French. 1988. "Dividend yields and expected stock returns." Journal of Financial Economics 22 (1): 3–25. ISSN: 0304405X. doi:10.1016/0304-405X(88)90020-7.

———. 2015. "A five-factor asset pricing model." Journal of Financial Economics. ISSN: 0304405X. doi:10.1016/j.jfineco.2014.10.010.

Folger, Jean. 2016. What is the relationship between inflation and interest rates? Accessed August 23, 2020. https://www.investopedia.com/ask/answers/12/inflation-interest-rate-relationship.asp.

Goyal, Amit. 2020. Amit Goyal - University of Lausanne. Accessed August 31, 2020. http://www.hec.unil.ch/agoyal/?fbclid=IwAR3Y4nnO5Lk7QPoBRTFsnQob8msArbvwbfHxj8weV-eP%7B%5C_%7DVn%7B%5C_%7DFrQpt9LEsIY%20http://www.hec.unil.ch/agoyal/.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." Econometrica 50 (4): 1029. ISSN: 00129682. doi:10.2307/1912775.

Investopedia. 2020. Risk-Free Return Definition. Accessed August 30, 2020. https://www.investopedia.com/terms/r/return.asp%20https://www.investopedia.com/terms/r/risk-freereturn.asp.

July Thomas and Samir Khan and Jimin Khim. 2019. Taylor Series Approximation — Brilliant Math & Science Wiki. Accessed August 31, 2020. https://brilliant.org/wiki/taylor-series-approximation/.

Macaluso, Jeff. 2018.
Testing Linear Regression Assumptions in Python - Jeff Macaluso. Accessed August 22, 2020. https://jeffmacaluso.github.io/post/LinearRegressionAssumptions/.

Maio, Paulo, and Pedro Santa-Clara. 2015. "Dividend Yields, Dividend Growth, and Return Predictability in the Cross Section of Stocks." Journal of Financial and Quantitative Analysis 50 (1-2): 33–60. ISSN: 17566916. doi:10.1017/S0022109015000058.

Monteiro, A N A Sofia. 2018. "Predictability of stock returns and dividend growth using dividend yields : An international approach CeBER Working Papers," no. 10.

Oslo Stock Exchange. 2020. Oslo Børs / Home - Oslo Børs. Accessed August 30, 2020. https://www.oslobors.no/ob%7B%5C_%7Deng/.

Sharpe, William F. 1964. "CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK." The Journal of Finance 19 (3): 425–442. ISSN: 15406261. doi:10.1111/j.1540-6261.1964.tb02865.x.

———. 1994. "The Sharpe Ratio." The Journal of Portfolio Management 21 (1): 49–58. ISSN: 0095-4918. doi:10.3905/jpm.1994.409501.

Shiller, Robert. 2019. Online Data - Robert Shiller. Accessed January 14, 2020. http://www.econ.yale.edu/~shiller/data.htm.

SSB. 2020.
09381: Pengemarkedsrente (NIBOR), styringsrenten, dagslånsrenten og bankenes rent
Accessed August 20, 2020.
https://www.ssb.no/statbank/table/09381/.

Thadewald, Thorsten, and Herbert Büning. 2007. "Jarque-Bera test and its competitors for testing normality - A power comparison." Journal of Applied Statistics 34 (1): 87–105. ISSN: 02664763. doi:10.1080/02664760600994539.

Tradingview. 2020. U.S. Indices — Quotes and Overview — TradingView. Accessed August 30, 2020. https://www.tradingview.com/markets/indices/quotes-us/.

Verdickt, Gertjan, Jan Annaert, and Marc Deloof. 2019. "Dividend growth and return predictability: A long-run re-examination of conventional wisdom." Journal of Empirical Finance. ISSN: 09275398. doi:10.1016/j.jempfin.2019.03.002.

Welch, Ivo, and Amit Goyal. 2008. "A comprehensive look at the empirical performance of equity premium prediction." Review of Financial Studies 21 (4): 1455–1508. ISSN: 08939454. doi:10.1093/rfs/hhm014. https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhm014.

Wohlner, Roger. 2020. Average Stock Market Return. Accessed August 31, 2020. https://www.wealthsimple.com/en-ca/learn/average-stock-market-return%20https://www.wealthsimple.com/en-gb/learn/average-stock-market-return.