Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

Kristian Øyen

Predictive Power in the U.S. Stock Market

An Empirical Analysis of the Effects of Parameter Uncertainty on Stock Market Predictions

Master's thesis in Financial Economy
Supervisor: Joakim Kvamvold
June 2023

NTNU
Norwegian University of
Science and Technology

Kristian Øyen

Predictive Power in the U.S. Stock Market

An Empirical Analysis of the Effects of Parameter Uncertainty on Stock Market Predictions

Master's thesis in Financial Economy
Supervisor: Joakim Kvamvold
June 2023

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

NTNU
Norwegian University of Science and Technology

FORORD

Masteroppgaven markerer slutten på fem års utdanning i et svært spennende felt. Jeg vil rette en stor takk til min veileder Joakim Kvamvold for gode tilbakemeldinger og nødvendig kritikk til oppgaven, samt stor dybdekunnskap i tolkning av resultater. Videre takker jeg fakultetet for oppfølging og bistand ved problemer med studiepoeng grunnet fagvalg, og en smidig løsning for å sørge for at masteroppgaven blir ferdig i ordinær tid. Til slutt takker jeg medstudenter Kristoffer Imset og Anders Udland for godt samarbeid gjennom graden.

ABSTRACT

This study analyzes the predictive ability of market variables on total returns in the S&P 500 Composite index and its sub-indices: Dividend Aristocrats, Growth, and Value. Valuation ratios and macro-factors for the market overall are examined as predictor variables on total returns for each index. The results show that dividend yield and Shiller's CAPE-ratio have significant predictive power on total returns, aligned with the Dividend Discount Model. The overall predictive power is low, indicating limitations in line with theory. Contrary to expectations, valuation ratios do not have a stronger predictor-coefficient for dividend companies, such as the Dividend Aristocrats index, than growth companies or the market overall. The Term Spread shows significant predictive power for the Value index, challenging previous research expectations. Non-stationarity in valuation ratios introduces bias and affects forecasting accuracy. These findings highlight the challenges in predicting index returns.

SAMMENDRAG

Denne studien analyserer prediksjonsevnen til markedsvariabler på totalavkastning i S&P 500 Composite-indeksen og dens underindekser: Dividend Aristocrats, Growth, og Value. Verdsettelsesforhold og makrofaktorer for markedet totalt sett undersøkes som prediktorvariabler på totalavkastningen for hver indeks. Resultatene viser at utbytteavkastning og Shillers CAPE-forhold har betydelig prediktiv kraft på totalavkastning, i tråd med Dividend Discount Model. Den generelle prediksjonskraften er lav, noe som indikerer begrensninger i tråd med teorien. I motsetning til hypoteser har ikke verdsettelsesforhold en sterkere prediktorkoeffisient for utbytteselskaper, slik som Dividend Aristocrats-indeksen, enn vekstselskaper eller markedet totalt sett. Rentespredning viser betydelig prediktiv kraft for Value-indeksen, og utfordrer tidligere forskning. Ikke-stasjonæritet i verdsettelsesforhold introduserer skjevhet og påvirker prognosenøyaktigheten. Disse funnene fremhever utfordringene med å forutsi indeksavkastning.

CONTENTS

CHAPTER

ONE

INTRODUCTION

I explore the predictive power of market variables on total returns in the S&P 500 Composite index and sub-segments. I use the Composite index as a proxy for the overall market. The sub-segments include the S&P Dividend Aristocrats, Growth and Value.

First, I replicate the study by Ang (2014) on predictive power in an Asset Management Model. I use a range of variables included in the literature of Ang (2014), Campbell and Shiller (2001) such as valuation ratios; dividend yield, earnings yield and Shiller's 10-year cyclically adjusted price-earnings ratio. I also include macro-factors such as government bond yields, inflation, and interest spreads. I standardize the variables in order to use the correlation coefficient to measure the amount of variance in the returns explained by the predictor variable (R^2). I run regressions using Newey-West standard errors, with the various predictors as explanatory variables, on log total return for the market. Total return differs from ordinary returns, since it includes dividends, interest payments and capital gains distributions. The total return is logged in order to sum the returns for long-horizon regressions with overlapping data.

I find significant predictors in dividend yield and Shiller's 10-year CAPE-ratio, consistent with the intuitions of the Dividend Discount Model by Gordon (1962). The predictive power is present, but low - measured by the correlation squared

(R^2). This is consistent with the R^2 limits calculated by Zhou (2010) and Ross (2005), confirming the low predictability in returns. The long-run predictive power is spurious (Ang, 2014). This means that the t-statistics prove a significant correlation, but the amount of predictive power is inflated, consistent with the limits of Zhou (2010).

Next, I extend the study to include sub-indices of S&P 500; Dividend Aristocrats, Growth and Value. I hypothesise that valuation ratios such as dividend yields have a higher predictive coefficient for dividend companies, specifically the Dividend Aristocrats index, than the Growth index and the market overall in the Composite index. I do not find this, which implies that dividend yield does not do a better job for dividend companies than growth companies and the market overall.

I find significant predictive power in the term spread for the Value index. The correlation is counter to the intuitions of Harvey (2011). A wider term spread should mean that the market outlook is positive, leading to higher stock returns. A negative correlation implies the opposite, that a wider spread means a positive outlook on the market. The comparison can't be made, since the study by Harvey (2011) is on market sentiment and not returns. Neither of us prove causality, and thus the finds should be interpreted with caution.

Finally, I extend the study to correct for stationarity. I argue that falling interest rates over 1996 to 2021 lead to extremes in the valuation ratios, similar to Campbell and Shiller (2001), but with a different cause. These extremes cause a structure break, which leads to non-stationary valuation ratios. This is inconsistent with theory, since it implies a linear increase in either dividends or price over time, which in turn leads to an increase or decrease in stock prices with a change in mean and variance over time. In the real data population, this does not hold in the long-run. Sooner or later, you see a downwards or upwards correction in the market, which is a reversion to mean, and an indicator of a stationary process. "Valuation ratios are random rather than deterministic" (Campbell & Shiller, 2001), meaning they can have biased coefficients in small samples. Forecasting is based on linear relationships, while the true relationship between valuation ratios and long-horizon returns might be non-linear, creating biased results.

CHAPTER

TWO

THEORY

I build on existing statistical models by Ang (2014) as the main framework, using existing theory on stock prices, dividends and earnings yields, cash flow positions and predictability in the variations of these factors. Authors in the field exclusively use aggregated data for the entire stock market. I also include sub-segments of the market indices such as Dividend Aristocrats, Growth and Value to add nuances to the existing literature. This is to see if there are any differences between the market sectors.

Specialists in the field believe dividend yield variation is the key to predicting returns, but have split opinions on which factors have predictive power. Campbell and Shiller (2001) show that dividend yields and earnings yields are important factors in forecasting stock performance. Cochrane (1992) believes expected returns is the key to forecasting variation in dividend yields, while Bansal and Yaron (2004) show the opposite - that cash flows are the sole forecaster. Ang's research falls somewhere in the middle, as he documents the ability to anticipate cash flows (Ang & Bekaert, 2007), while also recognizing that the volatility of dividend yields and equity returns change over time, implying that predictability in returns exists as well (Ang & Liu, 2007). Zhou (2010) calculates an upper bound of 5% for the R^2 of dividend yield and earnings yield on returns, meaning the explanatory power is severely limited.

Conventional theory on efficient markets states that the stock market is not predictable, meaning no valuation ratio such as the dividend-price and price-earnings ratio has any ability to forecast movements in stock prices (Fama, 1970). However, investors strive for excess returns. The main argument is that by carefully analyzing fundamental factors such as the valuation ratios mentioned above, or key metrics such as dividend yield and earnings, mispricing in the market can be identified and capitalized on to generate said excess returns. This is what sparks the debate on whether comprehensive market analysis can reliably yield returns greater than e.g. investing all your money into the S&P 500 Composite Index in the long-run (which is what an investor with high regard for the Efficient Market Hypothesis likely would do).

2.1 Main Framework

2.1.1 Dividend Discount Model

The Dividend Discount Model - also known as the Gordon Growth Model - states that the stock price, P, is the present value of future discounted dividends (Gordon, 1962):

$$P = \frac{D}{E(r) - g}, \tag{2.1}$$

where D is the expected dividend next period, $E(r)$ is the discount rate (expected returns), and g is the growth rate of dividends. When prices are high, either future expected returns are low, future dividend growth rate is high or a combination of both. You assume that when prices are high (comparing traditional valuation ratios), future growth is also high, since you pay for value in the future. However, forecasting shows that value stocks in general outperform growth stocks. This is the value effect, and results in high prices meaning low future growth - which is illogical based on investor intuition (Ang, 2014).

Mathematically rearranging Equation 2.1 results in:

$$E(r) = \frac{D}{P} + g \tag{2.2}$$

Equation 2.2 represents a common approach to estimating the expected return on a stock, where the expected return is equal to the sum of the dividend yield and the expected growth rate of cash flow. This means dividend yields should help forecast expected returns (Dow, 1920).

Based on the Dividend Discount Model rearrangement in Equation 2.2, returns are predicted by the dividend yield and the growth rate of cash flow. I expect this relationship to behave differently based on the payout structure of the company which earnings we are considering. Companies that regularly pay dividends and companies that focus on growth without paying dividends will have drastically different dividend yields, and the predictive power of the dividend yields on the earnings should then also be higher for the dividend-based "value" companies.

2.1.2 Predictability Regression

Theory says equity risk premiums are predictable, but the amount of predictability is small (Ang, 2014).

$$r_{t+1} = \alpha + \beta \times X_t + \epsilon_{t+1} \tag{2.3}$$

Equation 2.3 is a model that captures predictability in returns based on the relationship in Equation 2.2. r_{t+1} is the log returns for the next period and X_t is a set of predictive variables. The predictors range from market factors such as aggregate dividend yield and earnings yield, to macro factors such as bond yield, inflation, credit spread and oil prices (Ang, 2014). The model is a simple linear regression, with one predictor on the returns. The one-period short-run effect is in this sense the next 3 months. Equation 2.3 predicts the short-run effect on returns, while we are also interested in the long-run effects. Extending the model for multiple periods of return gives:

$$r_{t+4} + r_{t+3} + r_{t+2} + r_{t+1} = \alpha + \beta \times X_t + \epsilon_{t+4,4}, \tag{2.4}$$

which represents the next 12 months of returns.

Consider the residuals of Equation 2.4. The subscript of $\epsilon_{4,4}$ denotes that the residuals are realized at t+4, but involve returns over the last four periods. This is due to overlapping data - the same observations on returns are used in multiple returns sums. This overlap causes regular ordinary least squares inference to highly overstate the true predictability in the data (Ang, 2014).

2.1.3 Dividend Yield Variation

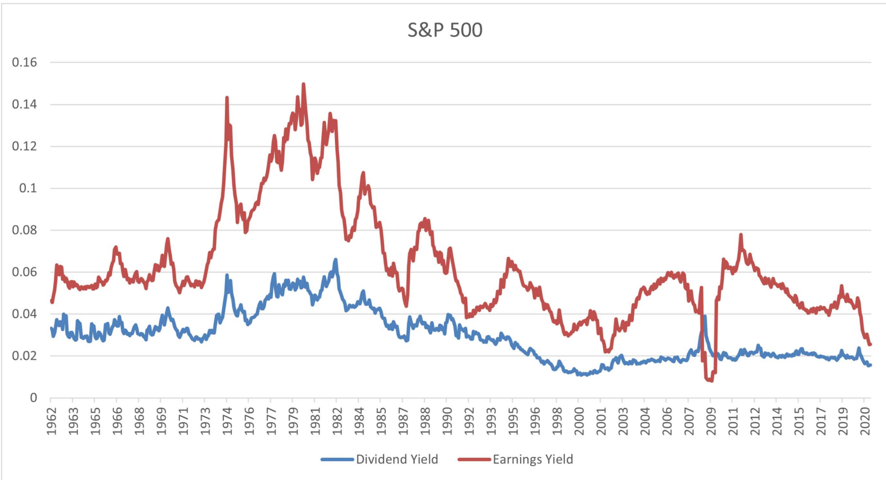

Figure 2.1.1: Dividend Yield & Earnings Yield for S&P 500 Composite

Ang (2014) argues that explaining dividend yield variation is the same as explaining what predicts returns. Both dividend yields and earnings yields move over time. Figure 2.1.1 shows both ratios from 1962 to 2021, where the correlation between the two ratios is 87%. They follow a similar trend - which is natural as the price is in the denominator for both ratios. In bad economic times, the yields tend to be high, and vice-versa. Note that during the financial crisis in 2008 the yields moved in opposite directions due to contracting earnings, and sticky dividend payout policies (Ang, 2014).

The question is whether the variation in dividend yield is moved by expected returns, cash flows or both. Cochrane (1992) states that all dividend yield variation comes from expected returns. Bansal and Yaron (2004) find that all dividend yield variation comes from cash flows. Ang and Bekaert (2007) document the existence of cash flow predictability, and show that dividend yields and equity return volatility vary over time. Based on this, the expected return must also be predictable.

Despite the discrepancies in the earlier work, the key takeaway is that dividend yields vary over time, and the variation is predictable through either expected returns, cash flows or a combination of both (Ang, 2014).

2.1.4 Parameter Uncertainty

The predictive coefficient - the β parameter in Equation 2.3 - varies over time (Ang, 2014). There is an inherent uncertainty surrounding the values and estimates that drive stock market dynamics, including expected returns, correlations and other fundamental characteristics. This uncertainty makes investment assessments difficult, and may cause the correlation relationships between asset classes to evolve over time. This can be accomodated by allowing coefficients to change slowly over time. Henkel et al. (2011) show regime-dependent predictive power. The predictive power is weak during recuperative periods, and strong during recessions, meaning the predictability is observed most during slow economic cycles.

2.2 How Much Predictability Can You Expect from an Asset Pricing Model?

Dividend yield and earnings yield have limited predictive power for future stock returns (Zhou, 2010). Zhou (2010) applies the predictability regression in Equation 2.3 with log returns as the dependent variable. He investigates the explanatory power of several valuation ratios, including dividend yield and earnings yield. They may provide some information about the attractiveness of stocks, but are not highly reliable predictors of future returns.

Zhou (2010) shows that the regression R^2 should be lower than 5%, arguing that the previous calculations of the same R^2 level found by Ross (2005) lacks specificity to be effective in real-world scenarios. He improves the limit by using a number that shows how closely related the state variables of an asset pricing model are to the default pricing kernel (Zhou, 2010). The pricing kernel is a concept used to determine the relative prices of different assets, that captures how investors make trade-offs between consuming today and saving for the future.

This means that we can expect at least 95% of the market movements to be unpredictable, and any claim to predict future returns with a higher R^2 than 5% should be viewed with great suspicion. The calculation of the limit accounts for overlapping data, which means that a longer horizon of overlapping data will lead to a greater bias in the explanatory power.

Grossman and Stiglitz (1980) show that profitable market-timing strategies are rare and statistically hard to detect, which coincides with the low predictive power found by Zhou (2010) and Ross (2005).

Ilmanen (2011) investigates the ability of market valuation ratios to predict future market returns. He shows that buying stocks when market valuations are cheap produce better returns than buying when valuations are rich. He uses correlations of predictors in line with the predictive regression of Equation 2.4, and finds that market timing is possible, but not easy. He emphasises that market timing is better during recessions, as return predictability are stronger due to the scarcity of risk-taking capital. Ilmanen (2022) is more reserved in his later papers, as he comments on limits of knowledge being a problem in efficient markets, causing limited return predictability. While the data available for analysis is increasing, we still have very limited data in the big picture.

2.3 Expanding on the Predictive Power of Valuation Ratios

Conventional valuation ratios such as dividend yield and CAPE-ratio are important forecasting variables for the U.S. Stock Market (Campbell & Shiller, 2001). Using Monte Carlo simulations for both dividend yield, price-earnings and stock prices that satisfy the efficient market hypothesis, Campbell and Shiller (2001) never obtain regressions that show as high forecasting ability as the actual data regressions. Their regressions should find low to no forecasting ability if the assumptions of the efficient market hypothesis holds, but they find forecasting ability in the data.

They explore valuation ratios used in forecasting stock prices. In short, they analyze stock prices, dividends and corporate earnings in the U.S. from 1871 to 1996. Stock prices tend to be more volatile than dividends or earnings, and dividend yields have historically been a more reliable predictor in forecasting stock returns than earnings metrics or price-earnings ratios. To elaborate on this, they use the requirement of the Efficient Markets Theory - that the dividend-price ratio (dividend yield) forecasts future dividend movements - and explore whether it instead forecasts future movements in stock prices.

2.3.1 Dividend Yield

Short-run dividend growth is predictable. Cash flow position and earnings play a huge role in the next quarters dividend policy, and combining this with historical payouts, the explanatory power is high. When using dividend yields to predict stock price changes over following years, Campbell and Shiller (2001) find low forecasting power. In times of historically high or low valuation ratios, you expect a reversion to the mean value at some point. The regression fits using the data reflects that this "restoration" is unpredictable. Campbell and Shiller (2001) argue that this is due to dividends not being the most accurate measure of fundamental value. This could sound strange, as dividends represents cash paid to shareholders.

Over long holding-periods, shareholder return is dominated by dividends, since the stock price return becomes small when it is discounted from the end of the beginning of a long holding period (Campbell & Shiller, 2001). Dividend yield can be affected by corporate payout policies, in the sense that companies can repurchase stocks, reducing the shares outstanding. This drives dividend growth up, and may permanently reduce the Dividend Yield ratio. This can be accounted for by looking at the "total shareholder yield" which combines dividend yield with the yield from share repurchases. Since dividend yield has low explanatory power on future dividend growth, they argue that the ratio forecasts movements in stock prices, which in turn reflects that it is the stock price (denominator) that reverts the ratio to the mean value over time.

2.3.2 Prices and Earnings

The Price-Earnings ratio adjusted for the 10-year moving average of real earnings has little predictability in forecasting earnings growth, both in the short- and long-run (Campbell & Shiller, 2001). The ratio is however a good forecaster of long-run growth in stock prices (ten-year horizon). Note that there are some econometric problems with the explanatory power in these long time spans.

As mentioned by Ang (2014), overlapping observations cause statistical dependence, inflating the explanatory power of the predictors. "Valuation ratios are random rather than deterministic" (Campbell & Shiller, 2001), meaning they can have biased coefficients in small samples. Forecasting is often based on linear relationships, while the true relationship between valuation ratios and long-horizon returns might be non-linear, creating biased results.

Data from historically normal economic times may not be a good predictor for economic times with abnormal valuation ratios and vice-versa, as the data itself is not comparable. Forecasting relations that worked in the past may not work in the future, due to societal changes and advancements in technology.

2.4 Yield Curve Inversion

A negative yield curve is a precursor of recession (Harvey, 2011).
The yield curve is the difference between interest rates with various time to maturities. Normally, the long-term rates are higher than the short-term rates, as you impose higher risk of macro-economic changes by locking your money for a longer time. Harvey (2011) finds a strong historic association with slow economic

growth or recession when the short-term rates that are higher than long-term rates.

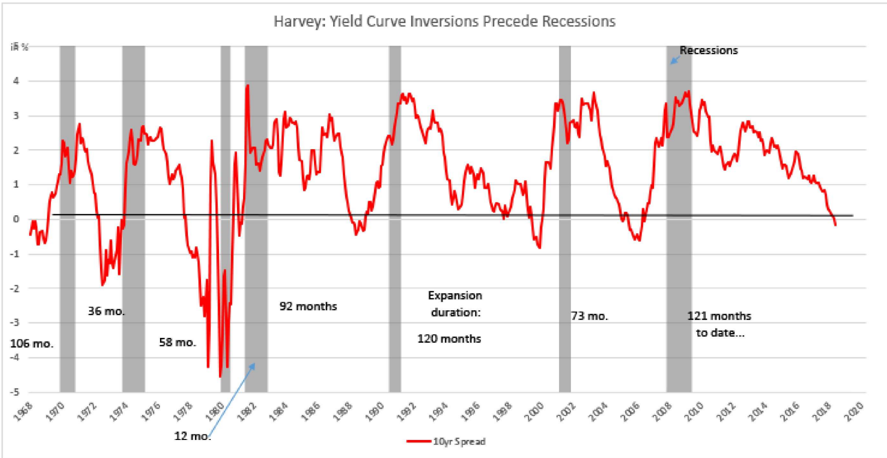

Figure 2.4.1: Harvey (2011): Yield Curve Inversions Precede Recessions. 10-Year Term Spread Paired with Economic Recessions

Figure 2.4.1 shows the 10-year Term Spread, which is the difference between 10-year government bonds and 3-month Treasury bill yields. Recessions are marked in grey. In front of each of the recessions since the 1960's, the yield curve inverts from positive to negative. Consider the U.S. 10-year Treasury bond. This is considered as one of the safest assets in the world, due to the economic power and recognition of the dollar. When market actors believe there is a risk of a crisis, they shift their economy into this bond. Buying pressure on the bond means that the prices rise, lowering the yields. If the long-term yield is lowered enough it ends up below the short-term yield, giving an inverted yield curve. The yield curve is thus an indicator of economic sentiment about the future (Harvey, 2011). The yield curve does not **cause** recession, but summarizes the sentiment of the economy. Since the inversion is a strong predictor of market sentiments, companies and consumers can use it to avoid additional risk in investments.

2.5 Significance of Dividend Yield for Dividend Companies

Dividend yields are usually higher in the Dividend Aristocrats index than the Growth index due to its focus on dividend-paying companies. Higher dividend yields attract investors that seek cash-flows, who may prioritize regular income generation over capital appreciation. Consequently, dividend yields influence the demand for stocks within the index. The Growth index consists of companies with high-growth potential. These companies typically reinvest profits into expanding operations, R&D, or acquisitions. This drives future growth, but as a result the dividend yields are lower and less consistent.

Companies must meet specific criteria to be included in the S&P Dividend Aristocrats, Growth, and Value indices. The Dividend Aristocrats index focuses on dividend-paying companies with a history of consistently increasing their dividends for at least 25 consecutive years. To be included, a company must be a member of the S&P 500 and meet liquidity and market capitalization requirements. In contrast, the Growth index consists of companies that exhibit high growth potential, typically characterized by strong earnings growth and revenue expansion. These companies are selected based on factors like sales growth, earnings growth, and price momentum. Lastly, the Value index includes companies that are considered undervalued relative to their intrinsic worth. The selection process for the Value index focuses on identifying companies with lower valuations and favorable financial fundamentals. These criteria ensure that the indices provide investors with diversified exposure to different market segments.

My hypothesis is that dividend yields in the market have a higher predictive power for returns in the Dividend Aristocrats index than the Growth index, since the payout policies are smoother. The dividend levels and growth in the Dividend Aristocrat index are stable, with diversified quality income that minimize risk.

The hypothesis can be formulated as:

$$H_0 : \rho_{DA} = \rho_G$$

$$H_\alpha : \rho_{DA} > \rho_G$$

where ρ_{DA}, ρ_G are the correlation coefficients between the dividend yields in the market and the respective returns on the Dividend Aristocrats and Growth indices.

For H_0 to be rejected, ρ_{DA} must be significantly different from ρ_G, measured by a significance test. This would mean that the dividend levels in the market are a better predictor for the Dividend Aristocrats index than the Growth index.

CHAPTER

THREE

DATA

I collect the dataset from Datastream (Refinitiv Eikon), Whartons Research Data Services (WRDS), and Federal Reserve Bank of St. Louis (FRED[1]). I don't have access to all stock variables needed for the study through WRDS, and therefore use Datastream for the missing variables. The data I collect is widely accessible through a range of sources, depending on accessibility. R. Shiller has calculated the CAPE-ratio for the S&P500 Composite Index that I have collected directly from his webpage through Yale University.[2]

3.1 Data Variables

Table 3.1.1 shows a summary of the variables I use. The prefixes show which index the data is collected from. r_{DA}, r_G, and r_V are the sub-index logged total returns for S&P500 Dividend Aristocrats, Growth, and Value respectively. Total return measures the change in the prices, dividends, interest and capital gains of the underlying index. r_M, Dividend Yield, Earnings Yield and Shiller P/E are market variables. I use the S&P500 Composite Index as a proxy for the stock market overall, as it is the best, and most accessible benchmark of U.S. Economy. Dividend Yield and Earnings Yield are the annual dividend and earnings yield, measured

[1]https://fred.stlouisfed.org/
[2]http://www.econ.yale.edu/ shiller/data.htm

Table 3.1.1: Descriptive Statistics of Measurement and Predictor Variables with Monthly Observations from January 1996 to January 2021. Values have been rounded to two/three decimals.

Variable	Obs	Mean	Std. dev.	Min	Max
r_{DA}	301	0.88%	4.03%	-14.69%	11.50%
r_G	301	0.83%	4.60%	-18.05%	13.50%
r_V	301	0.65%	4.66%	-18.76%	12.11%
r_M	301	0.75%	4.45%	-18.39%	12.06%
Dividend Yield	301	1.88%	0.37%	1.09%	3.89%
Earnings Yield	301	4.52%	1.26%	0.80%	7.79%
Shiller P/E	301	27.51	6.13	13.32	44.20
Long-Term Bonds	301	3.71%	1.57%	0.62%	6.90%
Short-Term Bonds	301	2.09%	2.02%	0.01%	6.17%
Inflation	301	2.14%	1.15%	-2.00%	5.50%
Term Spread	301	46.30%	37.47%	12.00%	335.00%
Credit Spread	301	99.21%	40.80%	55.00%	338.00%
Log Oil Price	301	0.37%	10.55%	-55.48%	46.91%

in percentages. Shiller P/E is the cyclically adjusted price-earnings ratio, using 10-year average earnings. I also include macro-factors. Long-Term Bonds is the yield on 10-year government bonds, as a measure on risk-free rate. I also include 3-month bonds, to capture the short-term interest structure. Inflation is change in consumer-goods prices, monthly observations for the change in the previous 12 months. Log Oil Price is the logarithmic change in dollar price for crude oil per barrel from the previous month. Term Spread is the 10-year treasury bill minus the 3-month treasury bill, reflecting the market's expectation for future interest rates and economic growth. Credit Spread is the absolute difference between low risk corporate bonds, and comparable higher risk corporate bonds (AAA-BAA). A higher spread generally means a higher perceived risk of default for the given bond. Companies with higher levels of debt, or industries that rely on credit are sensitive to the changes in these spreads.

The returns for each of the indices are what I am trying to predict. The reason I use total returns instead of excess returns is to include reinvestments of the earnings back into the companies. This considers the compounding effect of reinvesting dividends or interest. Total return is useful for the long-term predictability, since it provides a more accurate representation of the asset's performance where dividends or interest income contributes to the overall return. The valuation ratios are the most commonly used theoretical predictors, used by Ang (2014), Campbell and Shiller (2001), and Zhou (2010). The macro variables are also used by Ang (2014), to test for predictive power in interest rates, price changes and measures of economic growth.

Dividend and earnings yields both have a variation-coefficient[3] of 0.2. This should come as no surprise, as they are closely related, recall the high correlation explained in the theory. The returns and oil prices have variance-coefficients higher than one, which is made apparent due to the logarithmic transformation. This indicates skewed data observations.

3.2 Data Treatment

The 301 observations include monthly data from january 1996 to january 2021. This is long enough to capture any potential seasonal cycles. The sub-indexes were created as branches of the S&P500, so there is naturally far less data on the sub-indexes than the market overall. I only include data for the market in the time period where the sub-indexes exist to balance the data. This helps to generalize the data to a broader population.

3.2.1 Data Transformation

I use return variables for each of the sub-indexes, which is the current years Total Return Index measurement, divided by the previous year. I use this formula:

$$r_{t,i} = \frac{R_{t,i}}{R_{t-1,i}}, \tag{3.1}$$

where $r_{t,i}$ is the return for period t, for index i. $R_{t,i}$ is the total return index for index i for period t.

[3]Standard deviation divided by mean.

I then calculate log returns for each of the index returns, and sum the returns to obtain rolling quarterly, yearly and three-year return sums. By rolling, I mean monthly observations that capture the sum of the next 3, 12, 36 months of returns. This makes it possible to explore both the short- and long-run predictability. Log-returns are important, as you can only sum the returns for a given period after they are log-transformed.

3.3 Properties of the Valuation Ratios

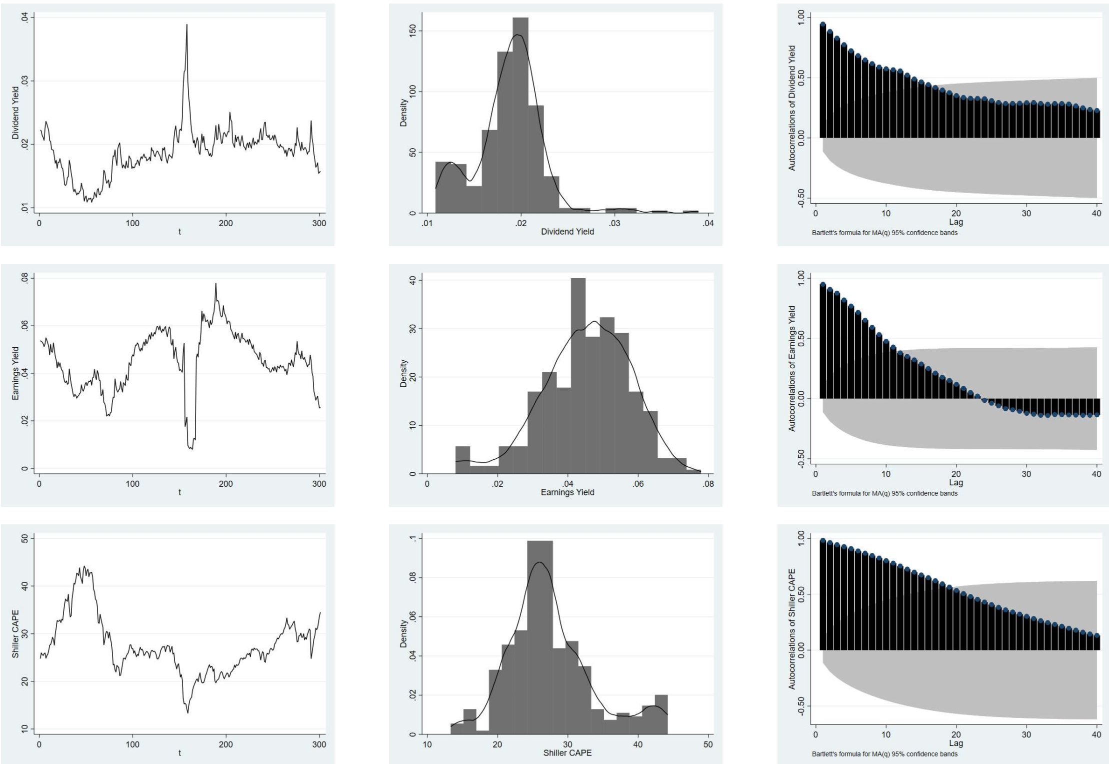

Figure 3.3.1: Time series properties of the dividend yields, earnings yields and Shiller PE, from top to bottom in that order. The left figures show line charts that plot the variables for each month. The middle figures show the density charts. The right charts show the autocorrelation functions.

Recall the positive correlation between dividend yield and earnings yield from Figure 2.1.1. Figure 3.3.1 show the same sudden spike in dividend yields against the drop in earnings yields (top left and middle left). As mentioned, this is due to the financial crisis in 2008. The density charts show the distributions of the ratios.

Most noteworthy are the right-hand charts, which show a positive autocorrelation for dividend yield and Shiller PE. Autocorrelation is the correlation between the observed value with a previous timestep of itself. In mathematical form this can be represented as $corr(y_t, y_{t-i})$ where i represents the number of previous timesteps (lags). The 95% confidence interval is marked by the gray area. Earnings yield is positively autocorrelated, and flattens around 23 lags, before becoming negatively autocorrelated.

For a stationary time series, the ACF will approach zero quickly with increasing lags. Looking at Figure 2.1.1, this is not the case for the valuation ratios, which is interesting. Consider the top right plot for dividend yields. Campbell and Shiller (2001) argue that either the price or dividend must revert the ratio to the mean over time, otherwise you would have a linear increase in prices or dividends, which as mentioned earlier is illogical in any real application. This should result in the ACF approaching zero very quickly, and the process being stationary. Instead, we see a flattening curve, which implies non-stationarity. The same can be said for earnings yield and Shiller P/E.

CHAPTER

FOUR

ANALYSIS

4.1 Standardized Variables

I standardize returns and the predictors. This is important, as the regression coefficient for a standardized variable represents the correlation between the predictive variable today, and the returns for the next period. (Ang, 2014).

Standardizing variables results in a z-score. I use the z-score when regressing the predictor variables on the returns. The formula for standardizing a variable X is given by:

$$z = \frac{(X - \bar{X})}{\sigma_X},$$

where $\bar{X}$ is the mean, and σ_X is the standard error.

This transforms the data to have a mean of zero, and standard deviation of one, without changing the underlying data structure. I interpret the z-score by looking at the sign of the value. A negative z-score indicates the observation is below the mean, while a positive z-score indicates the observation is above the mean. A z-score of zero means the data point is exactly at the mean. The size of a z-score, whether positive or negative, indicates the distance from the sample's standard deviation. This means we can use standardization to compare results across data that is measured in different ways, such as levels, percentages and differences.

After obtaining the z-score, I calculate the coefficient of determination for each of the standardized predictors on the returns:

$$R_X^2 = \rho_{Y,X}^2,$$

where R_X^2 is the coefficient of determination for the standardized predictor X, and $\rho_{Y,X}$ is the correlation-coefficient between the standardized returns Y and standardized predictors X.

By squaring the correlation coefficient, I obtain the proportion of the variance in the returns Y, that is explained by the predictor X. This makes it possible to investigate the explanatory power for each of the predictors on the returns using the correlation coefficient.

4.2 Newey-West Standard Errors

I use Ordinary Least Squares regression with Newey-West standard errors. This is a special case of OLS that adjusts the standard errors by considering the correlation structure of the data. The lagged values of the residuals and unequal variance is considered. This addresses two key issues in estimation. First, the Newey-West standard errors correct for serial correlation in the data variables. Second, they handle heteroskedasticity, which reduces the amount of inefficient estimates. This provides a more reliable inference when the data is assumed to be autocorrelated (Wooldridge, 2015). Small sample sizes cause Newey-West standard errors to be less efficient than regular Ordinary Least Squares, but over 300 observations is enough data to not lose information. I have already discussed variable selection and the functional form of the predictability regression model. Considering Newey-West corrects for heteroskedasticity and serial correlation, the statistical assumptions for reliable results hold true, and in this regard model specification is not an issue.

Newey-West estimation requires a selection of lag-length. In other words, how many previous time-steps of the variables to be included in the correction. If the lag-length is too strict, I might not account for serial-correlation adequately. If I include too many lags, I lose efficiency in the estimates due to overcorrection. I use a lag length of 4. There is no universal lag-length that applies to all situations, as it depends on the underlying structure of the time series. However, using a lag length of $T^{\frac{1}{4}}$, where T is the sample size, is fine for capturing all potential autocorrelation bias (Newey & West, 1987). See the mathematical derivation of this in Appendix B.1. It is important to note that the corrections when using Newey-West regressions may lead to larger standard errors compared to ordinary least squares estimation. This is a trade-off between efficiency and correction that is necessary in order to address bias in the data.

4.3 Regression Results

I run the regressions for the S&P 500 Composite index first, acquiring the t-statistics and correlation-coefficients for each of the predictors.

Table 4.3.1 reports correlation coefficients of the predictors on next-period returns at the beginning of the period. I use monthly observations of the predictors on monthly rolling quarterly, yearly and three-year sums of returns, indicated by the column titles "Q, 1Y, 3Y". Since I am using standardized variables, the coefficient of determination (R^2) of the predictor on the next-period returns is simply the correlation coefficient of said predictor squared. I have marked the significant coefficients in bold, the cutoff being 95% confidence - that is, a p-value smaller than 0.05. Focusing on dividend yield, we see a rising correlation for each of the columns, that is significant for the yearly and three-year sums of next-period returns, with an R^2 of 0.14 and 0.35 respectively[1]. The positive significant correlation is consistent with the Dividend Discount Model intuition in Equation 2.2, where high yields mean low prices that result from future cash flows being discounted at high expected returns. The Shiller CAPE Ratio give similar results, with significance in the yearly and three-year sums, with an R^2 of up to 0.29.

[1] 0.38^2 & 0.60^2

Table 4.3.1: Predictability Regressions of S&P 500 Composite with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021

Log Returns		Q	1Y	3Y
Dividend Yield	Correlation	0.19	**0.38**	**0.60**
	t-stat	(1.45)	**(3.74)**	**(5.06)**
Earnings Yield	Correlation	0.07	0.09	0.15
[1 Year Average Earnings]	t-stat	(0.67)	(0.75)	(1.18)
Shiller CAPE Ratio	Correlation	-0.10	**-0.24**	**-0.54**
[10 Year Average Earnings]	t-stat	(-0.92)	**(-2.33)**	**(-6.47)**
Government Bond Yield	Correlation	0.02	-0.006	-0.05
[10 Year]	t-stat	(0.23)	(-0.09)	(-0.63)
Government Bond Yield	Correlation	-0.04	-0.19	**-0.51**
[3 Month]	t-stat	(-0.55)	(-1.63)	**(-3.57)**
Inflation	Correlation	-0.04	**-0.13**	-0.07
[12-month Change in CPI]	t-stat	(-0.39)	**(-2.00)**	(-1.27)
Oil Price	Correlation	0.05	-0.10	-0.09
	t-stat	(0.49)	(-1.55)	(-1.70)
Credit Spread	Correlation	-0.18	-0.08)	-0.03
[AAA minus BAA]	t-stat	(-1.78)	(-1.11)	(-0.60)
Term Spread	Correlation	-0.23	-0.25	**-0.29**
[10 Year Treasury minus 3 Month T-bill]	t-stat	(-1.91)	(-1.43)	**(-2.15)**

Keeping in mind the R^2 limit calculated by Zhou (2010), this is too high - expecting a maximum of 0.05. The values are similar to what Ang (2014) find. The explanatory power is inflated due to overlapping observations, making them hard to justify. Newey-West underestimate the problem with overlapping observations, and thus overstate the predictability evidence (Ilmanen, 2022). The variables included are a range used in the literature; valuation ratios, macro factors and interest spreads. Newey-West regression accounts for the problem of time-varying volatility, which causes the regular ordinary least squares to overreject the null of predictability too often (Ang & Liu, 2007). Even with this correction, there is little to no evidence of predictability.

The Term Spread in Table 4.3.1 shows a rising determination coefficient (R^2) for all three returns sums, with the three year sum being significant. Harvey (2011) discusses the predictive power of the yield curve inversion on market sentiment, and market sentiment plays a big role in market returns, especially in a market period severely impacted by the falling interest rates. It's important to show restraint when interpreting the results since long-horizon returns are tricky, with econometric implications making the significant correlation values hard to defend (Ang, 2014).

Overall, the predictive power is low for most of the variables as expected. Earnings Yield, Long-term Bonds, Inflation, Log Oil Price and Credit Spread all have R^2 that comply with the limit of roughly 5% found by Zhou (2010) and Ross (2005). Short-Term Bonds only break the R^2 limit for the three-year sums, but this could also be due to the data structure which must be tested. Inflation is only significant for the yearly sums of returns, which is strange, since it is not significant for the three-year sums. I expected variables that are significant in one column to also be significant in the higher sums of years, since the overlap increases, inflating the predictive power. Inflation is not significant for the three-year sums. This could indicate a structure break in the data, or a similar bias. Short-term bonds are negatively significant for the three-year sum. This implies lower returns with increasing interest rates.

The overlapping returns in Equation 2.4 cause long-horizon R^2s to be inflated. The R^2 appears to indicate predictability when it isn't present. Consider the dividend yield correlation of 60% with log total returns over the next three years. This implies an R^2 of 35%, which is far larger than what theory by Zhou (2010) predicts. Long-horizon R^2s are spurious (Ang, 2014). The true predicability for the population is weak, but are much larger in small sample sizes. Ang (2014) refers to small sample sizes as hundreds or thousands of years. The overlapping observations cause false independence, meaning that the returns from t to $t+3$ overlap two periods with the period from $t+1$ to $t+4$. The longer the horizon, the worse the problem. The t-statistic corrections get the calculations right, but the R^2s are inflated. In essence this means that we can prove predictability to some extent, but how much predictability exists is hard to say, and theory suggests close to none.

When considering the newer sentiments of Ilmanen (2022), the low predictive power is in line with theory. Limited predictive power and limited data availability set heavy limitations on proving predictability in stock returns. There are continuously new challenges that need to be addressed, and it seems that we simply don't have enough knowledge of the inner workings of our equity markets to make any respectable claims of being able to predict it.

4.4 Extending to Sub-Indices of S&P 500

I explore how the results differ for the sub-indexes sorted by company category. This includes S&P Dividend Aristocrats, Growth and Value. The companies in these indices differ in their payout structure, and I mention briefly in the Theory chapter that I expect dividend yield to have a higher correlation with the earnings of companies structured on stable payout policies, than for companies who reinvest most of their earnings back into the company.

I run the regression using the same methods as for the Composite index, with Newey-West standard errors, reporting the correlation coefficients and t-statistics. Tables 4.4.1, 4.4.2 and 4.4.3 show the results. We see very similar results to the Composite index for all indices. Focusing on the index for Dividend Aristocrats in Table 4.4.1, the same parameters are significant, with some slight differences. The Term Spread is now significant for the quarterly sums instead of the three-year sums, and Inflation is no longer significant at any interval. Interestingly, we see a lower correlation for the dividend yield in this index than in the Composite index, suggesting that dividend yields are a weaker predictor for the Dividend Aristocrats index than the Composite index, which is the proxy for the market overall. Furthermore, if we look at Table 4.4.2, we see a **higher** correlation between dividend yields and next period returns for the Growth index than both the market and the Dividend Aristocrats index. To elaborate on what this implies; dividend yields in the market are a better indicator of future earnings for growth companies than dividend companies and the market overall. This suggests that dividend yields in the markets do not have higher predicitve power for earnings in the Dividend Aristocrats index than for earnings in the S&P Growth companies.

I test the difference in the correlation coefficients for dividend yields in Table 4.4.1 and 4.4.2 using a one-tailed t-test with a 5% significance level. The test-statistic will indicate whether we can reject the null hypothesis and accept the alternative. Recall the hypothesis of dividend yields in the market being a better predictor for returns in the Dividend Aristocrats index than the Growth index:

$$H_0 : \rho_{DA} = \rho_G$$

$$H_\alpha : \rho_{DA} > \rho_G$$

where ρ_{DA}, ρ_G are the correlation coefficients between the dividend yields in the market and the respective returns on the Dividend Aristocrats and Growth indices.

See Appendix C.1 for the one-tailed t-test results. Table C.1 reports the t-statistics for each time frame; quarterly, yearly and three-year sums of returns. The t-statistics are smaller than the critical t-value for all time frames, and we can not reject the null hypothesis. We can not say that dividend yields in the market are a better predictor of returns in the Dividend Aristocrats index than in the Growth index in this time period. This is not what I expected. One possible explanation for this is that while dividend yields are more stable and consistent for the Dividend Aristocrats index, I use total returns as the measure of performance. Total return includes reinvestments, which is typical for growth companies. The reinvestment portion of the earnings might account for the predictive power on the Growth index. Note that this is not something I have tested for, so the intuitions must be regarded with caution.

For the Value index in Table 4.4.3, the correlation between dividend yields and next-period returns is lower than any of the indices, implying lower predictive power for companies in this category. The Term Spread is now also significant for the quarterly sums. This is inconsistent with the intuitions by Harvey (2011).

For the negative correlation to be valid, an increasing Term Spread with higher long-term bond yields than short-term bond yields implies lower returns in the index total return. As the sentiment in the market of the future becomes more positive, leading to higher yields on long-run bonds since investors place their capital in stocks or short-term bonds instead, the returns become should increase according to Harvey (2011). Here, we see a decrease in the total returns as the long-run bond yields increase. Ang (2014) finds positive correlation between the Term Spread and log returns, consistent with Harvey (2011). I expect this is due to the difference in interest rates for the periods. Ang (2014) considers the period 1953 to 2011. From 1953 to 1981 the U.S. interest rates are steadily increasing, before decreasing until 2011. The period I am considering is from 1996 to 2021, with steadily decreasing interest rates for the whole period. This leads to lower bond yields, pumping the prices and valuation ratios to extremes, which in turn impacts the returns. There is also a notable difference in using total return as a measurement variable, than excess returns which could have implications on the correlation.

Table 4.4.1: Predictability Regressions on S&P Dividend Aristocrats with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021

Log Returns		Q	1Y	3Y
Dividend Yield	Correlation	0.15	**0.34**	**0.50**
	t-stat	(1.12)	**(3.08)**	**(8.34)**
Earnings Yield	Correlation	0.03	0.05	0.0007
[1 Year Average Earnings]	*t*-stat	(0.30)	(0.39)	(0.00)
Shiller CAPE Ratio	Correlation	-0.12	**-0.26**	**-0.46**
[10 Year Average Earnings]	*t*-stat	(-1.16)	**(-2.28)**	**(-6.49)**
Government Bond Yield	Correlation	0.02	-0.04	-0.04
[10 Year]	*t*-stat	(0.20)	(-0.53)	(-0.57)
Government Bond Yield	Correlation	-0.05	-0.18	**-0.57**
[3 Month]	*t*-stat	(-0.56)	(-1.58)	**(-4.83)**
Inflation	Correlation	-0.006	-0.12	-0.05
[12-month Change in CPI]	*t*-stat	(-0.04)	(-0.05)	(-0.72)
Oil Price	Correlation	0.05	-0.12	-0.04
	t-stat	(0.40)	(-1.67)	(-0.86)
Credit Spread	Correlation	-0.20	-0.08	-0.03
[AAA minus BAA]	*t*-stat	(-1.70)	(-0.96)	(-0.43)
Term Spread	Correlation	**-0.29**	-0.21	-0.18
[10 Year Treasury minus 3 Month T-bill]	*t*-stat	**(-2.83)**	(-1.23)	(-1.45)

Table 4.4.2: Predictability Regressions on S&P Growth with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021

Log Returns		Q	1Y	3Y
Dividend Yield	Correlation	0.24	**0.42**	**0.66**
	t-stat	(1.85)	**(3.37)**	**(4.85)**
Earnings Yield	Correlation	0.08	0.11	0.24
[1 Year Average Earnings]	t-stat	(0.78)	(0.98)	(1.88)
Shiller CAPE Ratio	Correlation	-0.12	-0.27	**-0.57**
[10 Year Average Earnings]	t-stat	(-0.99)	(-1.88)	**(-6.34)**
Government Bond Yield	Correlation	-0.006	-0.01	-0.03
[10 Year]	t-stat	(-0.08)	(-0.19)	(-0.45)
Government Bond Yield	Correlation	-0.08	-0.20	**-0.45**
[3 Month]	t-stat	(-0.74)	(-1.47)	**(-3.02)**
Inflation	Correlation	-0.09	**-0.13**	-0.07
[12-month Change in CPI]	t-stat	(-0.93)	**(-2.12)**	(-1.46)
Oil Price	Correlation	0.02	-0.11	-0.09
	t-stat	(0.16)	(-1.82)	(-1.74)
Credit Spread	Correlation	-0.11	-0.06	-0.02
[AAA minus BAA]	t-stat	(-1.25)	(-0.83)	(-0.31)
Term Spread	Correlation	-0.17	-0.21	-0.23
[10 Year Treasury minus 3 Month T-bill]	t-stat	(-1.45)	(-1.32)	(-1.79)

Table 4.4.3: Predictability Regressions on S&P Value with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021

Log Returns		Q	1Y	3Y
Dividend Yield	Correlation	0.12	**0.31**	**0.47**
	t-stat	(0.91)	**(3.25)**	**(5.24)**
Earnings Yield	Correlation	0.06	0.06	0.03
[1 Year Average Earnings]	*t*-stat	(0.56)	(0.52)	(0.26)
Shiller CAPE Ratio	Correlation	-0.07	**-0.21**	**-0.48**
[10 Year Average Earnings]	*t*-stat	(-0.70)	**(-2.38)**	**(-7.20)**
Government Bond Yield	Correlation	0.05	0.005	-0.06
[10 Year]	*t*-stat	(0.58)	(0.07)	(-0.79)
Government Bond Yield	Correlation	-0.01	-0.16	**-0.54**
[3 Month]	*t*-stat	-0.16	-1.64	**(-4.24)**
Inflation	Correlation	0.009	-0.11	-0.06
[12-month Change in CPI]	*t*-stat	(0.07)	(-1.77)	(-0.95)
Oil Price	Correlation	0.09	-0.07	-0.07
	t-stat	(0.79)	(-1.16)	(-1.35)
Credit Spread	Correlation	**-0.23**	-0.10	-0.05
[AAA minus BAA]	*t*-stat	**(-2.25)**	(-1.42)	(-0.96)
Term Spread	Correlation	**-0.28**	-0.27	**-0.33**
[10 Year Treasury minus 3 Month T-bill]	*t*-stat	**(-2.32)**	(-1.51)	**(-2.48)**

4.5 Correcting for Stationarity

Ang (2014), Campbell and Shiller (2001) do not correct for non-stationarity, but discuss the assumption of stationarity. I extend the study to include a correction for stationarity, since the valuation ratios are non-stationary. This is out of the ordinary, and a correction is interesting to see if it heavily changes the results.

The results in Table 4.3.1 differ from what Ang (2014) finds, looking at the bold t-statistics that indicate significance. He only finds significance for valuation ratios. Bond Yield, Inflation and Term Spreads are all insignificant in his regressions, while in Table 4.3.1, we see one of the sums being significant for each of the variables. This is not consistent with theory, and suggests a structure break in the data, causing seasonality or trends.

Ang (2014), Campbell and Shiller (2001) all discuss the forecasting ability of the valuation ratios, with special weight on dividend yield variation to hold predictive power on returns in the market. The ratios are relative valuation measures. *"The numerator or the denominator of the ratio must move in a direction that restores the ratio to a more normal level"* (Campbell & Shiller, 2001). In a case where they are non-stationary, they don't revert to mean over time. Consider non-stationarity in dividend yield. This implies a linear increase in either dividends or price over time, which in turn leads to an infinite increase or decrease in stock prices with a change in mean and variance over time. In real world applications, this does not hold in the long-run. Sooner or later, you see a downwards or upwards correction in the market, which is a reversion to mean, and an indicator of a stationary process.

The years in 1995-2021 is a period noted for an enormous market correction in 2008 during the global financial crisis, followed by a substantial rebound and market rally that lasted for several years. To aid this, the U.S. Federal Reserve Bank has steadily reduced the long-term interest rates. It is clear that there is a difference in this period to the periods considered by Ang (2014), Campbell and Shiller (2001). Figure 4.5.1 shows the 10-year government bond yield, approach-

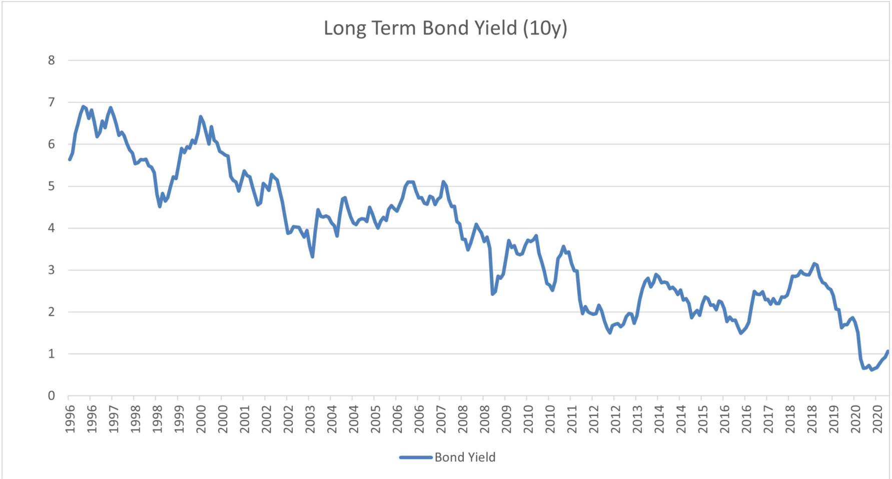

Figure 4.5.1: Long Term Bond Yield (10y) 1996 - 2021

ing values below 1% in 2020. This is around the same time when the yield curve inverts (Harvey, 2011). Lower interest rates lead to more money in the stock market, which steadily pumps up the valuation ratios to extreme values (Campbell & Shiller, 2001). Campbell and Shiller (2001) elaborate on whether the economic times of their study from 1872 to 2000 were replicable, due to the valuation ratios being so far from the historical averages, and the same can be said for 1996 to 2021.

4.5.1 Stationarity

I use the Augmented Dickey-Fuller test to check for stationarity in the data variables. Consider Equation 4.1:

$$\Delta X_t = \mu + \psi X_{t-1} + \sum_{i=1}^{i} \alpha_i \Delta X_{t-i} + u_t \tag{4.1}$$

The null hypothesis is that X is a random walk, with $\psi = 0$. In this case, the process is non-stationary.

A stationary process is one in which the probability distributions are stable over time. The time series is stationary if it has a constant mean, variance, and autocovariance regardless of when the observations are made (Wooldridge, 2015). If the process contains a unit root ($\psi = 0$), the value of the next observations will be random as there is no correlation with the observation in the previous time step.

In simpler terms, there is no upward or downward pressure to revert a sudden shock in the economy. Consider the effects of an economic collapse. If a unit root is present, the effects of the collapse will never subside, making the next time step in the process impossible to predict. A stationary process will eventually eliminate the shock effect, meaning there is a system that can be predicted to some extent.

4.5.2 Test Results

I test the time series variables before standardizing using the standard 5% significance level as the cut-off. See Appendix A.1 for the Augmented Dickey-Fuller test results. Testing for stationarity before and after standardization yields the same results, since standardization does not change the data structure, but scales the data for comparison purposes. Dividend Yield, Earnings Yield and Shiller's CAPE-ratio are all non-stationary in the given time-interval. The Term Spread is stationary, which means the variable does not need corrections to be included in the corrected regressions, and the results will be the same as for the Composite index.

4.5.3 Dealing With Non-Stationarity

I have non-stationary variables. To correct biased results in the regressions, I must first transform these variables. The variables that are non-stationary are first-differenced in order to remove trends.

Differencing a variable transforms it from representing the actual level of the variable, to the change in the observations between time periods. This effectively reduces or removes the trend component, because the level of the time series variable is removed (Wooldridge, 2015).

To illustrate this, I include Figure 4.5.2 that shows dividend yields as levels to the left, and as differences from the consecutive observations to the right. The spread is much more compact, suggesting that the trend element is removed. I discussed the Auto Correlation Function for the valuation ratios in the Data chapter that

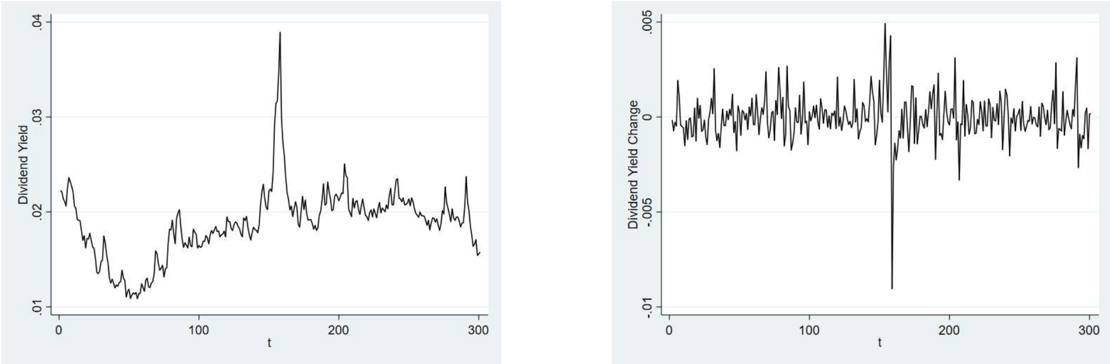

Figure 4.5.2: Dividend Yield vs. ΔDividend Yield

implied non-stationarity, which is confirmed by the ADF-tests.

It is important to note that this changes the data structure. We are no longer looking at how Dividend Yield itself predicts returns, but rather how the change in Dividend Yield from the previous month predicts returns. This is the case for all variables that are first-differenced. This is a simple method of correcting for non-stationarity. The point is not to perfectly correct all variables, but to reduce the data bias and see how the non-stationarity affects the results.

4.5.4 Stationary-Corrected Regressions

The final regressions for the Composite Index in Table 4.5.1 show that it is hard to find statistical evidence of predictability. The correlation coefficients and t-statistics are reported, and I highlight the significant results (5%) in bold.

The change in valuation ratios are not significant, suggesting that non-stationarity caused data bias for the valuation ratios on level format. Focusing on the coefficients, we see that they now all adhere to the limit calculated by Zhou (2010) and Ross (2005). The Term Spread for the 3-year return sums has a correlation of -0.29, which translates to an R^2 of 8.4%, which is the most extreme case. In other words, it is just as hard to predict returns as theory says it should be (Zhou, 2010; Ross, 2005). Inflation is only significant for the 1-year sums of returns. Furthermore, the R^2 is 1.7%, which is low. Interest rates and inflation are tightly connected, so it is likely that the falling interest rates over the period of consideration also play a big role in this.

Table 4.5.1: Corrected Predictability Regression on S&P 500 Composite with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021. Δ Indicates First-Difference

Log Returns		Q	1Y	3Y
ΔDividend Yield	Correlation	-0.09	-0.05	-0.003
	t-stat	(-1.16)	(-0.80)	(-0.05)
ΔEarnings Yield	Correlation	0.08	0.03	0.02
[1 Year Average Earnings]	*t*-stat	(1.85)	(0.56)	(0.50)
ΔShiller CAPE Ratio	Correlation	0.007	0.10	0.02
[10 Year Average Earnings]	*t*-stat	(0.08)	(1.26)	(0.29)
ΔGovernment Bond Yield	Correlation	0.02	-0.006	-0.04
[10 Year]	*t*-stat	(0.23)	(-0.09)	(-0.63)
ΔGovernment Bond Yield	Correlation	0.07	0.20	0.02
[3 Month]	*t*-stat	(0.65)	(1.40)	(0.20)
Inflation	Correlation	-0.04	**-0.13**	-0.07
[12-month Change in CPI]	*t*-stat	(-0.39)	**(-2.00)**	(-1.27)
Oil Price	Correlation	0.05	-0.10	-0.09
	t-stat	(0.49)	(-1.55)	(-1.70)
ΔCredit Spread	Correlation	-0.18	-0.08	-0.03
[AAA minus BAA]	*t*-stat	(-1.78)	(-1.10)	(-0.60)
Term Spread	Correlation	-0.23	-0.25	**-0.29**
[10 Year Treasury minus 3 Month T-bill]	*t*-stat	(-1.91)	(-1.43)	**(-2.15)**

4.5.5 Extending to Sub-Indices of S&P 500

I also run the regressions using the same right hand variables, swapping the returns of the S&P 500 Composite index for the sub-indices S&P Dividend Aristocrats, Growth and Value again, this time with the corrected market variables.

Table 4.5.2 show the correlations and t-statistics for the Dividend Aristocrats index. The results are very similar to the Composite index. Here, the change in 1-year averaged earnings yield is significant for the quarterly sums, which implies short-run predictability of returns for the period when using change in earnings in the market as the predictor. The R^2 is 0.8%, which again shows that the explanatory power is low, and it is hard to argue any practical use. The significance for change in earnings yield is consistent with Ang (2014), who also finds significance for the earnings yield. This was not the case before correction, which suggests that the non-stationarity inflate the predictive power. Again, it proves difficult to find statistical significance of predictability in returns.

Table 4.5.3 swaps the log returns with the Growth index, all else equal. The correlations are exceptionally low, for example the Shiller earnings ratio on quarterly sums of returns yielding -0.0003. This means the predictive power is effectively zero. It makes sense for the correlations to be lower for the Growth index. The companies are more focused on growing than steadily paying out cash, and as a result can have high returns one year and lower returns another year comparatively. Market conditions and earnings volatility play a significant role here. Investor sentiment is also likely a key part, especially with the falling interest rates, where focus on growth is severe. The high valuation ratios in the late 2010's to early 2020's show that the market expects high future returns. I write this knowing a severe correction is coming in 2022-2023, which reflects the cyclicity in the economy. The Value index, presented in Table 4.5.4, is similar to the Dividend index. The term spread is the most noteworthy variable, being significant for the quarterly sums and the three-year sums of returns. The Term Spread is as mentioned not corrected for stationarity due to not exhibiting non-stationary traits, so the sentiments will be the same as for the non-corrected regressions.

Table 4.5.2: Corrected Predictability Regressions on S&P Dividend Aristocrats with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021. Δ Indicates First-Difference

Log Returns		Q	1Y	3Y
ΔDividend Yield	Correlation	-0.06	-0.03	0.02
	t-stat	(-0.78)	(-0.45)	(0.27)
ΔEarnings Yield	Correlation	**0.09**	0.04	0.03
[1 Year Average Earnings]	*t*-stat	**(2.09)**	(0.57)	(0.54)
ΔShiller CAPE Ratio	Correlation	-0.07	-0.02	-0.08
[10 Year Average Earnings]	*t*-stat	(-0.72)	(-0.21)	(-1.58)
ΔGovernment Bond Yield	Correlation	0.02	-0.04	-0.04
[10 Year]	*t*-stat	(0.20)	(-0.53)	(-0.57)
ΔGovernment Bond Yield	Correlation	0.05	0.14	-0.07
[3 Month]	*t*-stat	(0.46)	(1.02)	(-0.82)
Inflation	Correlation	-0.01	-0.12	-0.05
[12-month Change in CPI]	*t*-stat	(-0.04)	(-1.95)	(-0.72)
Oil Price	Correlation	0.05	-0.12	-0.04
	t-stat	(0.40)	(-1.67)	(-0.86)
ΔCredit Spread	Correlation	-0.21	-0.08	-0.03
[AAA minus BAA]	*t*-stat	(-1.7)	(-0.96)	(-0.43)
Term Spread	Correlation	**-0.29**	-0.21	-0.18
[10 Year Treasury minus 3 Month T-bill]	*t*-stat	**(-2.83)**	(-1.23)	(-1.45)

Table 4.5.3: Corrected Predictability Regressions on S&P Growth with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021. Δ Indicates First-Difference

Log Returns		Q	1Y	3Y
ΔDividend Yield	Correlation	-0.06	-0.04	0.003
	t-stat	(-0.97)	(-0.69)	(0.06)
ΔEarnings Yield	Correlation	0.05	0.01	0.008
[1 Year Average Earnings]	*t*-stat	(1.26)	(0.21)	(0.22)
ΔShiller CAPE Ratio	Correlation	-0.0003	0.10	0.02
[10 Year Average Earnings]	*t*-stat	(-0.003)	(1.37)	(0.31)
ΔGovernment Bond Yield	Correlation	-0.006	-0.01	-0.03
[10 Year]	*t*-stat	(-0.08)	(-0.19)	(-0.45)
ΔGovernment Bond Yield	Correlation	0.01	0.11	0.02
[3 Month]	*t*-stat	(0.09)	(0.86)	(0.23)
Inflation	Correlation	-0.09	**-0.13**	-0.07
[12-month Change in CPI]	*t*-stat	(-0.93)	**(-2.12)**	(-1.46)
Oil Price	Correlation	0.02	-0.11	-0.09
	t-stat	(0.16)	(-1.82)	(-1.74)
ΔCredit Spread	Correlation	-0.11	-0.06	-0.02
[AAA minus BAA]	*t*-stat	(-1.25)	(-0.83)	(-0.31)
Term Spread	Correlation	-0.17	-0.21	-0.23
[10 Year Treasury minus 3 Month T-bill]	*t*-stat	(-1.45)	(-1.32)	(-1.79)

Table 4.5.4: Corrected Predictability Regressions on S&P Value with Newey West Standard Errors - Various Market Predictors on Log Total Return as Dependent Variable for Quarterly, Yearly & Three-Year Sums of Monthly Returns from January 1996 to January 2021. Δ Indicates First-Difference

Log Returns		Q	1Y	3Y
ΔDividend Yield	Correlation	-0.10	-0.06	-0.008
	t-stat	(-1.28)	(-0.89)	(-0.156)
ΔEarnings Yield	Correlation	**0.10**	0.05	0.03
[1 Year Average Earnings]	t-stat	**2.29**	(0.89)	(0.79)
ΔShiller CAPE Ratio	Correlation	0.02	0.08	0.009
[10 Year Average Earnings]	t-stat	(0.19)	(1.09)	(0.17)
ΔGovernment Bond Yield	Correlation	0.05	0.005	-0.06
[10 Year]	t-stat	(0.58)	(0.07)	(-0.79)
ΔGovernment Bond Yield	Correlation	0.13	0.28	0.01
[3 Month]	t-stat	(1.28)	(1.93)	(0.14)
Inflation	Correlation	0.009	-0.11	-0.06
[12-month Change in CPI]	t-stat	(0.07)	(-1.77)	(-0.95)
Oil Price	Correlation	0.09	-0.07	-0.07
	t-stat	(0.79)	(-1.16)	(-1.35)
ΔCredit Spread	Correlation	**-0.23**	-0.10	-0.05
[AAA minus BAA]	t-stat	**(-2.25)**	(-1.42)	(-0.96)
Term Spread	Correlation	**-0.28**	-0.27	**-0.33**
[10 Year Treasury minus 3 Month T-bill]	t-stat	**(-2.32)**	(-1.51)	**(-2.48)**

CHAPTER

FIVE

CONCLUSION

The data analysis shows that there is low predictability in index returns, indicating the difficulty of accurately forecasting returns. The results align with the research conducted by Ang (2014) on the S&P Composite index. when extending the analysis to sub-indices for S&P; Dividend Aristocrats, Growth, and Value, the results remain consistent. My hypothesis of dividend yields in the market doing a better job of predicting total returns in the Dividend Aristocrats sub-index than the Growth sub-index and the market proxy did not hold in the one-tailed t-test, indicating that dividend yields do not have a higher predictive power on the Dividend Aristocrats index than the Growth index and the market overall.

The negative term spread correlation is counter to the intuitions of Harvey (2011). A wider term spread should mean that the market outlook is positive, leading to higher stock returns. A negative correlation implies the opposite. I do not prove causality, but reflect whether predictive power exists or not, based on the t-statistics. In short - the results don't prove what causes the relationship, just that the predictive power is present. Furthermore, said predictive power is low, so the results should be regarded with caution.

Extending the study to correct for non-stationarity fortifies the indications of low predictive power, as the R^2 limit of 5% calculated by Zhou (2010) holds for an array of predictor variables used in the literature. Decreasing interest rates play a role in the structure of the data, as valuation ratios are pumped to an extreme in the time period. This is paired with the non-stationarity in the valuation ratios that cause bias in the regressions.

I use total return as the measure of performance to include earnings reinvested, which is relevant for the sub-indices I study. An extension to this thesis could explore the relationship with excess returns, which is more common in the literature, to see if the hypothesis results are the same. You could also test other variables such as the earnings yield to see if the predictive power is significantly better for one of the indices than the rest, although arguing what practical use this will have is hard, since the predictive power itself is so low. In general economists agree on the low predictive power of valuation ratios, while recessions seem to be the best period to execute investment strategies based on predictions. Nothing in this study indicates any predictive power that can be used in an investment strategy.

In summary, this analysis demonstrates the challenges in predicting index returns and suggests no significant difference in the predictive power of dividend yields on dividend stocks compared to growth stocks and the overall market. The observed negative correlation between the term spread and stock returns challenges conventional expectations, and caution should be exercised in interpreting these findings due to the complexities of the underlying factors and the presence of non-stationarity. Predictability is present, but too low to be effectively used in any investment strategies. Even the term spread that Harvey (2011) argues is an effective predictor of recessions is not a solid predictor on returns, but rather on market sentiments.

REFERENCES

Ang, A. (2014). *Asset management: A systematic approach to factor investing.* Oxford University Press.

Campbell, J. Y., & Shiller, R. J. (2001). Valuation ratios and the long-run stock market outlook: An update.

Gordon, M. J. (1962). The savings investment and valuation of a corporation. *The Review of Economics and Statistics*, 37–51.

Zhou, G. (2010). How much stock return predictability can we expect from an asset pricing model? *Economics Letters*, *108*(2), 184–186.

Ross, S. A. (2005). *Neoclassical finance.* Princeton university press.

Harvey, C. R. (2011). Yield curve inversions and future economic growth. *Diakses dari https://faculty. fuqua. duke. edu/˜ charvey/Research/Professional_ Materials/Term_ Structure_ May_ 17_ 2011. pdf.*

Cochrane, J. H. (1992). Explaining the variance of price–dividend ratios. *The Review of Financial Studies*, *5*(2), 243–280.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, *59*(4), 1481–1509.

Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, *20*(3), 651–707.

Ang, A., & Liu, J. (2007). Risk, return, and dividends. *Journal of Financial Economics*, *85*(1), 1–38.

Fama, E. F. (1970). Efficient market hypothesis: A review of theory and empirical work. *Journal of Finance*, *25*(2), 28–30.

Dow, C. H. (1920). *Scientific stock speculation.* Magazine of Wall Street.

Henkel, S. J., Martin, J. S., & Nardari, F. (2011). Time-varying short-horizon predictability. *Journal of financial economics*, *99*(3), 560–580.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393–408.

Ilmanen, A. (2011). *Expected returns: An investor's guide to harvesting market rewards* (Vol. 535). John Wiley & Sons.

Ilmanen, A. (2022). *Investing amid low expected returns: Making the most when markets offer the least*. John Wiley & Sons.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 705.

APPENDICES

APPENDIX

A

AUGMENTED DICKEY FULLER TEST RESULTS

A.1 ADF test

(ADF Critical values; 5% = -2.86 1% = -3.43)

Variable	**t-adf**	**Δt-adf**
r_M	-16.401	
r_{DA}	-16.453	
r_G	-16.642	
r_V	-16.024	
Dividend Yield	-2.836	-16.095
Earnings Yield	-2.447	-18.771
Shiller CAPE	-1.265	-14.660
Long Term Bonds	-0.881	-13.784
Short Term Bonds	-1.128	-11.380
Inflation	-3.184	
Term Spread	-4.513	
Credit Spread	-2.573	-10.687
Oil Price	-13.528	

Table A.1: Table of ADF-statistics on regression variables[1]

[1]Note: If the process is non-stationary, the variable is differenced once, to correct for non-stationarity. All variables in this set are stationary at a maximum of one difference.

APPENDIX

B

LAG LENGTH SELECTION NEWEY-WEST STANDARD ERRORS

B.1 Mathematical Derivation of Lag Length

Let m be lag-length, growing with the sample size T, and allow T to assume all positive integers. Newey and West (1987) show that their estimator for the co-variance matrix is consistent if the lag length satisfies two conditions:

1. Lag length m grows with the sample size T.

$$\lim_{T\to\infty} m(T) = \infty,$$

2. Lag length m grows at a slower rate than $T^{\frac{1}{4}}$.

$$\lim_{T\to\infty} [\frac{m(T)}{T^{\frac{1}{4}}}] = 0,$$

Considering the above, the lag length can in all reasonable econometric sense be set to the integer of $T^{\frac{1}{4}}$.

Using a sample size of 301, the lag length is 4:

$$m^* = 301^{\frac{1}{4}} = 4.16$$
$$m = 4$$

APPENDIX

C

HYPOTHESIS TEST RESULTS

C.1 One-Tailed T-Test

Table C.1: Hypothesis Test Results

Time Frame	Difference	Standard Error	t-Statistic
Quarter	-0.09	0.1811	-0.496
Year	-0.08	0.1626	-0.492
Three-Year	-0.16	0.1466	-1.092

df	600
Significance level	0.05
Critical t-value	1.645

Table C.1 shows the difference between the correlation coefficient of dividend yields in the market with the earnings of the Dividend Aristocrats index and the Growth index:

$$\Delta_\theta = \rho_{DA} - \rho_G$$

The standard error is calculated from the standard errors in the Newey-West regressions for each of the indices:

$$\sigma_\theta = \sqrt{\sigma_{DA}^2 + \sigma_G^2}$$

The t-statistic is then:

$$t = \frac{\Delta_\theta}{\sigma_\theta}$$

This process is repeated for each of the time frames.

NTNU

Norwegian University of
Science and Technology