Harald Blichfeldt Bjerke
Robin Larsen

Economic Activity, Inflation and Yield Equilibrium:

A Deep Dive into Stock-Bond Dynamics

Master's thesis in Financial Economics
Supervisor: Svein-Arne Persson
June 2023

Master's thesis

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

NTNU
Norwegian University of
Science and Technology

Harald Blichfeldt Bjerke
Robin Larsen

Economic Activity, Inflation and Yield Equilibrium:

A Deep Dive into Stock-Bond Dynamics

Master's thesis in Financial Economics
Supervisor: Svein-Arne Persson
June 2023

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

NTNU
Norwegian University of
Science and Technology

Abstract

Over the past decades, the negative correlation between returns on bonds and equities has become a cornerstone of diversified portfolios. Recent developments have seen historic drawdowns in 60/40 portfolios, so understanding the causal mechanisms is of interest to every market participant who is restricted to long-only allocations in the two asset classes.

The primary objective of this research is to study under which economic conditions positive and negative correlations of returns between risk-free bonds and risky equities can occur. By using time series data, the effect of both the activity level and the inflation rate on the yields of stocks and bonds are estimated. The earnings yield of the S&P 500 Index, and the 10-Year U.S. Treasury yield is studied. The quarterly data stretches back 60 years, thus covering a wide range of economic developments.

The findings reveal a risk-adjusted equilibrium between the yields of these two assets and show that these yields exhibit causal relationships with each other. In terms of economic conditions, periods of inflationary pressure tend to trigger negative returns in both asset types. In contrast, under normal conditions, the stage is set for regimes where bond and equity returns are negatively correlated.

Sammendrag

Over de siste tiårene har den negative korrelasjonen mellom avkastningen på obligasjoner og aksjer blitt en hjørnestein i diversifiserte porteføljer. Nylige hendelser har gitt historisk store tap i 60/40-porteføljer, så å forstå de kausale mekanismene er av interesse for enhver markedsdeltaker som er begrenset til kun lange allokeringer i de to aktivaklassene.

Hovedmålet med denne oppgaven er å studere under hvilke økonomiske forhold positive og negative korrelasjoner av avkastning mellom risikofrie obligasjoner og risikable aksjer kan forekomme. Ved å bruke tidsseriedata, blir effekten av både aktivitetsnivået og inflasjonsraten på avkastningen av aksjer og obligasjoner estimert. "Yieldene" til S&P 500-indeksen og den 10-årige amerikanske statsobligasjonen blir betraktet. De kvartalsvise dataene strekker seg tilbake 60 år, og utgjør dermed et bredt spekter av økonomiske utviklinger.

Resultatene tyder på en risikojustert likevekt mellom "yieldene" på disse to aktivaklassene og de har kausale effekter seg imellom. Når det gjelder økonomiske forhold, har perioder med inflasjonspress en tendens til å utløse negative avkastninger i begge aktivaklasser. På den andre siden, under normale forhold, er scenen satt for regimer der obligasjoner og aksjeavkastning er negativt korrelert.

Acknowledgements

This Master's thesis concludes our Master of Science in Financial Economics at the Norwegian University of Science and Technology. We extend our gratitude to our supervisor, Svein-Arne Persson, and our dedicated reader, Birger Vikøren, for their valuable guidance in times of great need. A sincere thank you to Solveig too, for her contributions and encouragement.

Contents

1 Introduction

This paper explores the potential existence of an equilibrium between the yields of the S&P 500 index and 10-Year U.S. Treasuries, considering unemployment and inflation as state variables, from 1959 to 2018. The primary goal of this paper is to study effects of inflation and economic activity on returns for these two assets. Specifically, under what conditions are the correlation between them negative, and when are they positive? In our analysis, we shed light on the short and long-term mechanisms inherent to diversified equities and risk-free treasuries, their causality, and their sensitivities to economic activity and inflation. Understanding these mechanisms are of interest to academics, policymakers, and investors, particularly those constrained to long-only allocations in these two assets.

The inspiration for this analysis stems from a public investment research study conducted by Johnson et al. (2013), which proposes the existence of regimes exhibiting positive correlation of returns between the two assets during inflationary pressures. However, their paper does not fully address the inherent relationship between stocks and bonds. As such, we also draw inspiration from Zakamulin and Hunnes (2019), who explore the causality between the two assets. The contribution of our analysis lies in its comprehensive consideration of the stock-bond causality, and macroeconomic effects.

We formulate dynamic Error Correction models that allow us to estimate long term relationships in the two yields. Our framework builds on what is known as the Fed Model, which unrealistically assumes the absence of a risk premium. The disadvantage of our modeling choice is that we cannot reject this hypothesis, but the advantage is that we obtain reliable results concerning the short term dynamics of returns in the two assets, which is the primary goal of this paper.

In conclusion, our findings align with those of the aforementioned studies for the most part. We identify evidence of a long-term equilibrium in the yields of the two assets, and describe the causality between them. Moreover, we find a long-term relationship between the two yields and our macroeconomic state variables. When evaluating long- and short-term effects, our results also support the existence of positive correlation regimes in inflationary environments. However, diagnostic results indicate a lack of robustness, and weaknesses that are common to the study of financial returns.

The paper is organized as follows: Chapter 2 delves into the yields, their inherent relationship, and the general effects of inflation and unemployment. Chapter 3 explains econometric terms and empirical tests later used. Chapter 4 presents other approaches and results from related literature before detailing our empirical specification for the analysis. Chapter 5 provides information on our data and includes a graphical examination. The empirical analysis begins in Chapter 6, starting with tests for stationarity, followed by tests for cointegrating relationships and estimation of our Error Correction models. Diagnostic tests are also conducted. A summary and conclusion is provided in 7.

2 Theory

This chapter will explain the two most central concepts for our thesis, the earnings yield and the treasury yield. We show how their first differences approximate returns for their respective assets. Importantly, we explain the theoretical effects of inflation and unemployment on the two yields and their returns. We also explain the causal channels between equities and treasuries.

2.1 Earnings Yield

The earnings yield is a widely used financial ratio and estimate of expected return that expresses a company's earnings per share as a percentage of its share price (Ilmanen, 2011). Generally, the yield can be defined as

$$\text{Earnings Yield}_t = \frac{E_t}{P_t} \tag{2.1}$$

where P_t denotes the market price of the asset at time t, and E_t denotes the earnings of either an equity or market index.

The earnings yield can be expressed in several ways, mostly depending on how the earnings component is measured. E_t can represent forward looking estimates from analysts, the last twelve months (LTM) of earnings, or even a long time frame average. The idea is to obtain the best estimate of future earnings, thus allowing the expected return to be comparable to asset classes other than equities.

We define our earnings yield as an average of the last ten years of earnings:

$$\frac{\bar{E}_t}{P_t} = \frac{\frac{1}{10}(E_{t-1} + E_{t-2} + \cdots + E_{t-10})}{P_t} \tag{2.2}$$

Because earnings are reported quarterly, 40 quarters will make up $\bar{E}_t$:

$$\frac{\bar{E}_t}{P_t} = \frac{\frac{1}{40}(E_{t-1} + E_{t-2} + \cdots + E_{t-40})}{P_t} \tag{2.3}$$

This definition is based on inverting the Cyclically Adjusted Price-to-Earnings ratio, also known as the CAPE ratio or Shiller P/E, made famous by Campbell and Shiller (2001). Their research introduced a new take on the widely used Price-to-Earnings (P/E Ratio) multiple which measures what an investor pays for the LTM of earnings and thus whether a stock is overvalued or undervalued.

In the context of the P/E multiple, Campbell & Shiller normalized earnings by using average 10-year earnings, adjusted for inflation. This ratio is defined as

$$\text{CAPE}_t = \frac{P_t}{\frac{1}{10}(E_{t-1} + E_{t-2} + \cdots + E_{t-10})} \tag{2.4}$$

where P_t refers to the inflation-adjusted spot price at time t, and $E_t - 1$, $E_t - 2$ etc. refers to the lagged, inflation-adjusted earnings. The same inflation adjustment holds for the earnings yield in (2.2).

The disadvantage of using LTM earnings or forward estimates in the calculation of the earnings yield is that they can be highly sensitive to cyclicality and volatility from one-off cost and revenue increases. As such, this can occasionally result in deceivingly high or low observations of the earnings yield (Ilmanen (2011)). This is especially true when evaluating individual companies that are prone to idiosyncrasies, as opposed to diversified indices. These observations hold true for valuation ratios in a general sense. Less variation in E_t is advantageous in the context of this paper, so averaging the earnings seems like a better idea than using LTM or forward estimates[1]. The CAPE ratio has also shown promising predictive power.

Angelini et al. (2012), found that this definition of the yield was a good estimator for long-run returns. These findings were substantiated by Jivraj and Shiller (2017), who found that this definition was the most consistent predictor amongst common and similar yields[2]. Research focusing on the CAPE ratio´s predictability (Klement, 2012) have also found similar results.

Much of the criticism regarding CAPE, and by extension the earnings yield (2.2), has revolved around its ability to revert to historical mean, especially in

[1] The approach is also supported by Ilmanen (2011).
[2] The other yields: NIPA/P, B/P, S/P, E/P, D/P, CF/P.

the past 20 years. According to Shiller, the CAPE was never intended as a market timing tool that would be expected to revert to its historical mean after brief periods of high or low values. Shiller recognizes its historical tendency to mean-revert, but states that it might be more interesting as an estimator for prospective returns of equities in the context of asset allocation or relative valuation (Bunn and Shiller, 2014; Jivraj and Shiller, 2017). This was also highlighted in Ilmanen (2011), which classified CAPE as a relative value indicator.

The CAPE ratio has received criticism that naturally relates to the yield as well, more recently from Siegel (2016). Siegel acknowledges CAPE's historical accuracy, but argues that fundamental changes to corporate payout policy might have reduced its accuracy considering share repurchases have become increasingly preferred over dividends. Share repurchases have been beneficial for both investors and companies, with the former avoiding the dividend tax and the latter being able to reduce outstanding shares and by extension its P/E ratio. Ultimately, this change in corporate payout policy could result in an increase in per share earnings (EPS) growth rate, thus introducing an upward bias on the level of the CAPE ratio (downward bias on the CAPE yield).

Furthermore, Grantham (2017), also argues that EPS growth may be one of the reasons for a downwards bias in the CAPE ratio, but points to increased profitability amongst many of the S&P 500 companies as one of the drivers. According to him, there could be several fundamental reasons for this, hereunder greater monopolistic power, globalization, falling interest rates and higher leverage.

2.1.1 Earnings yield and returns

The earnings yield has a convenient connection to returns in the underlying asset. It can be shown that its change is an approximation of returns. To illustrate, consider the change in $\bar{E}_t$ from one quarter to the next:

$$\Delta(\frac{\bar{E}_t}{P_t}) = \frac{\frac{1}{10}(E_{t-1} + E_{t-2} + \cdots + E_{t-40})}{P_t} - \frac{\frac{1}{10}(E_{t-2} + E_{t-3} + \cdots + E_{t-41})}{P_{t-1}} \quad (2.5)$$

where Δ is the first difference operator.

From one quarter to the next, the price P_t can change substantially. For the earnings component however, the inner 38 quarterly earnings that make up $\bar{E}_{t-1}$ will also be part of $\bar{E}_t$. This means most of the variation in $\Delta(\frac{\bar{E}_t}{P_t})$ will be due to changes in P_t, by a large order of magnitude. As such, $\Delta(\frac{\bar{E}_t}{P_t})$ can be seen an approximation of returns, albeit inversely.

2.2 Treasury Yield

We define the 10-Year Treasury Yield as the risk free rate. It represents the return on investment for a risk free security. We symbolize the Treasury Yield as

$$\text{Treasury Yield}_t = Y_t$$

A treasury note is risk free if it is held to maturity. If the note is not held to maturity, the note holder is exposed to interest rate risk. As such, market pricing can cause negative or positive returns from time t to $t+1$ even in risk free treasuries.

The issuance of these treasuries follow an auction-based process where anyone can bid, from large private investors to corporations, institutions and governments. This is called the primary bond market. The bidders specify the rate, yield or discount rate they will accept, which determines the price of the bond. (Thau, 2010)

The price of a coupon-bearing bond, such as the U.S. Treasury Notes, can be generalized to equation (2.6) (McDonald, 2013). The definition follows the concept of present value of future cash flows. Coupon payments and principal repayment is discounted using a discount rate, i.e. the yield, reflecting the market's required rate of return for holding the bond. From this process and these factors, the Treasury's yield to maturity, Y_t, can be derived, representing the annualized rate of return earned by an investor assuming the bond is held to maturity[3] (Thau, 2010). The price of a coupon bearing bond, as a function of the yield, can be written as:

$$B(y) = \sum_{i=1}^{n} \frac{C/m}{(1+y/m)^i} + \frac{M}{(1+y/m)^n} \tag{2.6}$$

Here, $B(y)$ is the bond price with m coupon payments per year for T years. Per-period coupon payments is denoted C/m, while M represents the Face Value

[3]The yield can later change as the Treasuries starts trading in the secondary market. However, we will only be considering treasuries on a constant maturity basis throughout this paper.

paid at maturity. y/m is the per-period yield to maturity, and $n = mT$ is the number of periods until maturity.

2.2.1 Treasury yield and returns

The inverse relationship between the prices of risk free treasuries and their yields form the interest rate risk. When yields increase, the discount rate applied to the present value of future coupon payments also rises, causing bond prices to decrease, and vice versa.

The relationship, however, is not linear but rather convex due to the effect of compounding present in the discount factor in the denominator. To exemplify this, consider equation (2.6). Any change in yield (y) results in an increasingly pronounced effect on $B(y)$ as i and n increase. This means that fluctuations in the yield will have a more significant impact on the price of bonds with longer maturities (Berk and DeMarzo, 2017).

Duration serves as a conventional linear approximation of the convex sensitivity between bond prices and yields[4]. When duration is higher, it implies that the bond's cash flows are received further into the future, making the bond more sensitive to changes in interest rates. So likewise, bonds with higher duration will experience larger price fluctuations in response to changes in yields compared to bonds with shorter duration (Thau, 2010; Berk and DeMarzo, 2017).

Moving forward, we will use an approximation for the price sensitivity of a coupon bearing bond with respect to its yield, as given by Berk and DeMarzo, 2017. In this formula:

$$r^y \approx -D_{Mod} * \epsilon \tag{2.7}$$

In this equation, r^y represents the return on the bond, D_{Mod} is the modified duration, and ϵ is a small change in yield, y.

So, if we rearrange (2.7), we can express changes in the risk-free rate, ΔY, in terms of the bond return per unit of duration:

$$\Delta Y = -r^y \tag{2.8}$$

[4]Duration and convexity are both measures of the sensitivity of bond prices to changes in yield. Duration is calculated as the value-weighted average maturity of a bond's cash flows and provides a linear approximation of this sensitivity. Convexity, on the other hand, provides a more accurate measure by taking into account the curved relationship between bond prices and yields, which duration does not capture. The difference in accuracy between duration and convexity tends to be minor (Thau, 2010).

This shows that changes in the risk-free rate are negatively related to the bond return, adjusted for its duration. This approximation is more accurate for minor changes in Y. The error increases exponentially with the change in Y.

Let's now consider the impact of time on this relationship by reintroducing the temporal dimension. When the duration of the bond changes only slightly from time t to $t+1$, which is typically the case when T is large, the approximation remains valid. A bond's duration will decrease with time, all else equal. Therefore, in the short term, the approximation holds true for treasuries with a long time to maturity:

$$\Delta Y_t = -r_t^y \tag{2.9}$$

2.3 Macroeconomic effects

Macroeconomic conditions form market participants' expectations for risk and profits. In this subchapter, we explain how inflation and economic activity affect the earnings yield, the treasury yield, and importantly their respective returns.

2.3.1 Unemployment

The unemployment serves as an important proxy of economic activity and output. With high economic activity, production of goods and services is usually also high (Begg et al., 2014). As the activity level increases, more employees are needed in order to increase production and unemployment decreases. Conversely, decreasing economic activity generally increases unemployment. Hence, their relationship is inverse.

Equities and treasury yields are both influenced by unemployment levels and general economic activity. When unemployment is high, or rising, reduced consumer spending and lower corporate profits tend to follow, often leading investors to anticipate future economic downturns. This uncertainty may lower expectations of future corporate earnings, inciting a shift from riskier equities to safer assets like treasuries. Conversely, during times of low or falling unemployment, consumer spending and corporate profitability typically increase, encouraging a transition from safe assets to riskier equities.

This mechanism closely resembles the consumption-based theory of Cochrane (2005), arguing that higher expectations for future consumption increases the agent's willingness to invest today, substituting away from risk free saving alternatives.

Thus, one can expect falling unemployment to induce lower treasury yields (higher prices/returns). Conversely, rising unemployment induces higher earnings yield (lower prices/returns).

2.3.2 Inflation

According to Ilmanen (2011) and Neville et al. (2021), inflation generally has a negative effect on equity prices, especially short term inflation changes. Neville et al. (2021) points to multiple causes. First of all, revenues and operating margins may decrease due to reduced purchase power among customers, and increased costs on raw materials and wages, respectively. Secondly, abrupt increases in inflation may cause uncertainty among both investors and company leaders regarding the overall state of the economy. Consequently, investments could be reduced as both parties seek to reduce their risk. For companies, this could also harm long-term growth and thus their returns.

Another popular suggestion as to inflation's negative impact on equities is the popular inflation-illusion (or money-illusion) hypothesis, suggested by Modigliani and Cohn (1979). This hypothesis points to the possibility of investors and analysts incorrectly using nominal discount rates rather than real rates when valuing equities. As such, periods of increasing inflation would also increase discount rates, leading to irrationally low valuations from falling equity prices. The opposite would be true for periods of declining inflation. (Ilmanen, 2011)

As for bonds, an increase in inflation is generally considered to have a negative impact due to its embedded inflation expectation, according to Neville et al. (2021). As such, rising inflation would generally result in higher bond yields. This is especially true for bonds with short maturity and low duration. The impact on bonds with higher duration, however, would depend more on whether the increase in inflation is perceived to result in a more persistently high inflation level. If this is the case, high duration bonds would be more sensitive to inflation increases than low duration bonds.

To summarize, rising inflation is generally expected to cause negative returns both in equities and risk free treasuries. Likewise, higher inflation increases the earnings yield and the treasury yield.

2.4 The relationship between stocks and bonds

In this subchapter, we establish and explain two channels of causality between risky equities and risk free treasuries themselves - "The PV channel", and "The substitution channel". They are intended to explain how the causality between the treasury yield and the earnings yield can vary.

2.4.1 "The PV channel"

This channel is uni-directional, going from the treasury yield, Y_t, to the earnings yield, $\bar{E}_t/P_t$.

Recall that Y_t represents the risk-free rate. Investors evaluate the value of risky equities by discounting their future expected cash flows to the present value. If the risk-free rate increases, it raises the discount rate, thereby reducing the present value of future cash flows, and thus the price of the equities, all else being equal. This results in an increase in the earnings yield given that earnings yield is inversely related to the price of equities.

Additionally, a higher risk-free rate can increase the cost of capital for businesses. This could potentially reduce their profitability, further lowering the price of the equities.

Therefore, a higher risk-free rate often leads to lower equity prices, negative returns on equities and consequently higher earnings yields. This mechanism is a probable source of one-way causality, going from the treasury yield to the earnings yield. We term this "The PV Channel" for ease of reference.

2.4.2 "The substitution channel"

This channel describes a positive and bi-directional relationship between the treasury yield and the earnings yield, and can be understood as a substitution effect. When the yield of one asset increases, investors may move their capital away from the unchanged asset towards the first asset, thereby increasing the yield of that second asset.

Consider an investor who allocates capital between a risk-free security (treasury notes) and risky equities. If the risk-free rate increases, it becomes more attractive, prompting the investor to reallocate capital by selling equities. As the price of equities falls (from the selling pressure), the earnings yield (E/P) increases, because earnings yield is inversely related to the price of equities.

Conversely, if the expected return on equities, as represented by the earnings yield, increases, it can attract more investment away from risk-free securities. As investors sell treasury notes to buy equities, the price of the notes falls, causing their yield to increase.

In this way, the yields of risk-free securities and equities are mutually dependent and can exert influence on each other due to substitution effects as investors rebalance their portfolios in response to changes in expected returns.

2.5 Fed Model, and extension

In this subchapter, we start by explaining what is known as the contentious Fed Model. We explain its unrealistic assumptions, and postulate that it can hold true on a risk adjusted basis.

The Fed Model is given by

$$\frac{E_t}{P_t} = Y_t \tag{2.10}$$

The equation implies that in the long term, the yield on risky equities should be equal to that of risk-free treasury notes, implicitly assuming a risk premium of 0. According to the Fed Model, equities are considered more attractive than treasuries when the ratio of expected earnings (E_t) to price (P_t) exceeds the yield on treasuries (Y_t), and vice versa. When this equality holds true, both assets are deemed fairly valued.

The Fed Model can be aligned with financial theory, but only under unrealistic assumptions. As such, it has faced criticism from scholars such as Asness (2002), Micaletti (2020), Estrada (2006) and Arnott et al. (2018). To illustrate this point, we can consider the Gordon Growth Model, which calculates the intrinsic value of a stock:

$$P_t = \frac{D_t(1 + k_t)}{r_t^f + RP_t - k_t} \tag{2.11}$$

In this equation, P_t denotes the price, D_t the dividend, r_t^f the risk free rate, k_t the dividend growth, and RP_t the risk premium.

Estrada (2006), shows that both sides of the equation are divided by earnings E_t when the following assumptions are made:

- all earnings are distributed as dividends, $D_t(1+k_t) = E_t$,
- dividends do not grow, $k_t = 0$,
- and investors demand no excess return for equities over treasuries, $RP_t = 0$

Under these assumptions, (2.11) simplifies to:

$$\frac{P_t}{E_t} = \frac{1}{r_t^f} \implies \frac{E_t}{P_t} = r_t^f = Y_t \tag{2.12}$$

and we obtain the Fed Model. The assumption of $RP_t = 0$ is unrealistic and fundamentally contradictory of most financial theory.

Typically, an excess return is required by investors, and the earnings yield should exceed the treasury yield as compensation. The assumption of $RP = 0$ can be relaxed to allow for a time-varying and non-zero risk premium. In that case, (2.11) simplifies to:

$$\frac{P_t}{E_t} = \frac{1}{r_t + RP_t} \implies \frac{E_t}{P_t} = Y_t + RP_t \tag{2.13}$$

This equation describes an equilibrium in the two yields on a risk adjusted basis. However, two considerations arise. First, the risk premium is unobservable. Second, factors such as uncertainty and macroeconomic shocks shape expectations for future earnings and risks.

Thus we postulate that the risk adjusted equilibrium is determined by macroeconomic state variables. The economic activity, g_t, affects corporate profits, and by extension the demand for risk-free securities. The inflation rate, inf_t, affects expectations for future risk-free rates. Together they would capture a lot of the variation in the risk premium. Hence, we propose a more encompassing equilibrium condition to the original Fed Model:

$$\frac{E_t}{P_t} = Y_t + inf_t + g_t \tag{2.14}$$

We postulate that this equation holds true in the long term. In an efficient market, a risk adjusted equilibrium should exist between the yields of these two assets. Any deviation from this equilibrium could provide savvy investors with a potential arbitrage opportunity, or at least, superior risk-adjusted returns. By seizing these opportunities, the equilibrium would be restored.

Summary

In this chapter, we have provided the theoretical background for the economics of our analysis. We established a bidirectional relationship between the earnings and treasury yields, through "The Substitution channel". Another one-way channel of causality from the risk free rate to equities was also explained, which we called "The PV channel".

Two equilibria were established. One parsimonious where no excess return is required, equation (2.12). The second one is more specific, in that it allows for a positive risk premium determined by the state variables inflation and economic activity, equation (2.14). We also showed how the yields in question can be approximated to their respective returns in the short-run.

Furthermore, the expected effects of inflation and economic activity on returns in the two assets were established. As explained, higher economic activity induces positive returns for equities, but the opposite for treasuries, while inflation is expected to cause negative returns in both assets.

The following chapter presents a general and compact explanation of important properties of time series. After, Chapter 4 includes an empirical specification based on the theoretical background presented in Chapters 2 and 3.

3 Time series properties

This chapter explains the fundamental properties of time series, the concept of stationarity and how to formally test for it. The chapter explains the concept of cointegration and links it to Error Correction Models. After, we explain diagnostic tests that will be used in the estimation of our models.

3.1 Stationarity

Every variable in this paper are time series. They must meet the criteria for ordinary least squares (OLS) estimation to give the best unbiased linear estimates (BLUE) (Wooldridge, 2015). Specifically, the series must be stationary.

To qualify as weakly stationary, a time series must generally maintain a constant mean, constant variance, and a time-independent autocovariance (Brooks, 2008), formally expressed as:

$$E(y_t) = E(y_{t-s}) = \mu \tag{3.1}$$

$$Var(y_t) = Var(y_{t-s}) = \sigma^2 \tag{3.2}$$

$$Cov(y_t, y_{t-s}) = Cov(y_{t-j}, y_{t-j-s}) = \gamma_s \tag{3.3}$$

If either of these conditions are not met, the process is considered non-stationary. Alternatively, the process can be strictly stationary if the joint distribution of any subset of the time series values remains the same when the whole series is shifted in time. Formally, for all times $t, t+1, ..., T$ and for all s, the joint distribution of the variables $(y_t, y_{t+1}, ..., y_T)$ is the same as the joint distribution of $(y_{t+s}, y_{t+1+s}, ..., y_{T+s})$:

$$f(y_t, y_{t+1}, ..., y_T) = f(y_{t+s}, y_{t+1+s}, ..., y_{T+s}) \tag{3.4}$$

This means that the entire probability distribution of the series must be time-invariant. In the context of this paper, weak stationarity is sufficient for the analysis, and strict stationarity is uncommon for the kind of real world data that we consider. Hence, the rest of this paper will consider weak stationarity.

A general first order autoregressive (AR1) process often serves as a helpful starting point for exploring weak stationarity. This is a linear combination of the last period's previous value of the dependent variable (Brooks, 2008):

$$y_t = \beta y_{t-1} + \mu + u_t \tag{3.5}$$

μ denotes a constant, u_t is a stochastic residual term assumed to be white noise, and β captures the predictive power of the lagged value of y. Given this formulation, one can calculate the mean and variance for an entire series within a sample. Provided $\beta < 1$, equation (3.5) satisfies the requirements for weak stationarity and consequently displays stability and a long-term equilibrium solution. This can be shown by the following steps:

If we substitute in for y_{t-1}, y_{t-2} and so on, (3.5) can be rewritten as a Moving Average (MA) process (Brooks, 2008). This process is a linear combination of the white noise residuals, where y_0 denotes the starting value:

$$y_t = \mu \sum_{i=0}^{t-1} \beta^i + \beta^t y_0 + \sum_{i=0}^{t-1} \beta^i u_{t-i} \tag{3.6}$$

As time goes on, (3.6) becomes:

$$\lim_{t \to \infty} y_t = \frac{\mu}{1-\beta} + \sum_{i=0}^{\infty} \beta^i u_{t-i} + \beta^t y_0 \tag{3.7}$$

Thus, if $\beta < 1$ and t is large, the final term of (3.7) converges to zero. u_t has expectation 0, and would similarly converge to 0 in the case of shocks. Thus, the equation reduces to the long term mean of

$$E(y_t) = \frac{\mu}{1-\beta} \tag{3.8}$$

This means that any shocks that create deviations from said long term mean will be corrected with time. This behavior is referred to as "mean reversion".

3.1.1 Non-stationarity

The typical non-stationary process is a random walk, which can be expressed as:

$$y_t = y_{t-1} + \mu + u_t \tag{3.9}$$

In this equation, nearly identical to (3.5), the coefficient in front of y_{t-1} is 1. $\beta = 1$ implies that y has a unit root, leading to a non-stationary process. The stationarity conditions, constant mean, variance, and covariance (equations 3.1 to 3.3), are hence violated. The process follows a random walk. When a shock occurs, its impact will persist over time, distinguishing it from a stationary process where the effect of a shock dissipates. When β exceeds 1, the influence of a shock will escalate over time and the series diverges and "explodes".

The constant μ in (3.9) also plays a significant role in determining the behavior of the process. When $\mu > 0$, it causes the growth of y to follow a deterministic trend. On the other hand, when $\mu = 0$, the process becomes a random walk without drift given $\beta = 1$. In such a scenario, even in the absence of a trend, the mean and variance will continue to expand with time.

Random walk processes can generally be differenced once in order to obtain a stationary series. This means that the process contains a unit root, and is integrated of order 1, denoted $I(1)$. Generally, y_t is stationary after the dth difference, $I(d)$.

Another reason why a time series may be non-stationary is when the series contains a structural break. This means that the population function of the time series changes over the sample period. A structural break can change the mean, variance and autocovariance of the series (Brooks, 2008). However, before and after the structural break occurs, the series will be stationary.

3.2 Identifying stationarity

Testing for stationarity is done by testing for unit roots (Dickey and Fuller, 1979). We want to test

$$y_t = \beta y_{t-1} + \mu + u_t$$

for $\beta = 1$ under the null, against the alternative $\beta < 1$. By subtracting y_{t-1}, we obtain

$$\begin{aligned}\Delta y_t =& \mu(\beta - 1)y_{t-1} + u_t \\ =& \mu + \psi y_{t-1} + u_t\end{aligned}$$

where $\psi = \beta - 1$ and Δ is the first difference operator.

The Augmented Dickey-Fuller (ADF) test is only valid in the abscence of serial correlation (Brooks, 2008). u_t is assumed to be white noise, and so it cannot be serially correlated, but it will be if Δy_t exhibits serial correlation. Thus, we add q lags of the first difference to control for serial correlation in Δy_t. The selection of q is an empirical question. We obtain:

$$\Delta y_t = \mu_y + \psi y_{t-1} + \sum_{i=1}^{q} \pi_1 \Delta y_{t-i} + u_t \tag{3.10}$$

Based on equation (3.10), we can now test $\beta = 1$ and $\psi = 0$ simultaneously:

H_0: $\beta = 1$, $\psi = 0$, the time series is non-stationary and contains at least one unit root.

H_1: $\beta < 1$, $\psi < 0$, the time series is stationary and contains no unit roots.

The test statistic for the ADF is obtained with a t-test:

$$\text{Test statistic} = \frac{\hat{\psi}}{sd(\hat{\psi})} \tag{3.11}$$

where $\hat{\psi}$ is the OLS estimate while $S\hat{E}(\psi)$ is the estimated standard error of $\hat{\psi}$.

The test statistic does not follow the standard t-distribution under an ADF test, but rather a non-standard distribution that depends on the model formulation of the alternative hypothesis. That is, to what extent it contains a constant and/or a trend (Brooks, 2008). This means that we must use the more stringent Dickey Fuller critical values when testing. The null hypothesis is rejected if the test statistic is more negative than the critical value.

Several approaches are available in the selection of lags, q, when testing (3.10). One such is the use of Information Criterion (IC) that calculates the following tradeoff: Too many lags reduce the degrees of freedom in the ADF test and reduce the ADF test statistic. Too few lags will not sufficiently control for serial correlation in Δy_t and bias the results (Brooks, 2008).

Akaice's Information Criterion (AIC) is given by

$$AIC = \ln(\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2 + \frac{2k}{T}) \tag{3.12}$$

where T is the number of observations, $k = q+p+1$ is the number of lags q and variables p. The first part of the IC measures the goodness of fit by considering the residual sum of squares. The second term penalizes excessive lags, q, and explanatory variables, p. Generally, we select the model with the lowest IC.

3.3 Testing for cointegration

The first requirement for the existence of cointegration between two variables is that they must be integrated of the same order. If two variables are non-stationary but integrated of the first order, they will cointegrate if there is a linear combination of the two variables that is stationary. Assuming a static equilibrium relationship between two variables, y and x, this can be written as:

$$y_t = \mu + \beta x_t + u_t \tag{3.13}$$

This equation is also called the cointegration equation. μ is a constant and β is a scaling parameter. We can test (3.13) for cointegration by studying u_t which can be interpreted as the deviation from equilibrium. Given cointegration, u_t will have expectation 0 and vary around this level. Further interpretation can be done by assuming that the deviation follows an AR(1) process. Its first difference can be written as

$$\Delta u_t = \theta u_{t-1} + v_t \tag{3.14}$$

where v_t is white noise. By testing whether the residual u_t is stationary, i.e., $\theta < 0$, one can examine whether y_t and x_t cointegrate. Stationarity in u_t is the second requirement for cointegration in equation (3.13).

The original cointegration test, from Engle and Granger (1987), involves the following procedure. First, the static cointegration equation given by (3.13) is estimated using the least squares method. Then the residuals which represent the differences between the observed and predicted values—essentially deviations from the equilibrium, are saved:

$$\hat{u}_t = y_t - \hat{\mu} - \hat{\beta} x_t \tag{3.15}$$

Next, it is examined whether the residuals are stationary using the ADF test described earlier. Rejecting the null hypothesis that $\hat{u}_t$ is non-stationary would mean that there exists an equilibrium relationship between y_t and x_t that holds in the long run. Given $0 < \theta_t < 1$, a positive deviation in the equilibrium will be corrected through the growth rate Δu_t in equation (3.14).

3.4 Error correction models

If the two variables form a cointegrated relationship, we can formulate a dynamic error correction model (ECM) according to Engle and Granger (1987). Given cointegration in equation (3.13), every variable in the ECM will be stationary, meaning inference can be done in a regular fashion.

The ECM captures how the long term equilibrium of y and x is restored through the growth rate of y: Δy is a function of changes in x, a constant term, and the error correction term, which captures the deviation from the long-term equilibrium in the previous period:

$$\Delta y_t = -\alpha z_{t-1} + \pi \Delta x_t + \gamma + v_t \tag{3.16}$$

Δy_t is the first difference of y. π captures the short run elasticity of y w.r.t. x. Any long term effect of x on y will be captured by the coefficient β in the cointegration equation (3.13). γ denotes the constant term and z_{t-1}, is generally known as the error correction term. It is a rewriting of $\hat{u}_t$ from (3.15), and captures any deviation from the long-term equilibrium in the previous period.

Given cointegration in (3.13), $0 < \alpha < 1$ (Enders, 2014). It captures the speed of adjustment towards the long-term equilibrium, i.e., the proportion of the initial deviation from equilibrium that is eliminated in the subsequent period. By n, periods, α^n parts of the initial deviation will have been corrected.

The Granger representation theorem states that if the variables in (3.13) are cointegrated, there exists an ECM, and vice versa (Enders, 2014). Thus, α in equation (3.16) should be significantly different from 0 if there is cointegration. That means a t-test of $\alpha = 0$ tests for an error correction mechanism in (3.16) and cointegration in (3.13) simultaneously, by definition. This can be a helpful validation of the cointegration tests described earlier.

3.5 Diagnostic tests

In this subsection we explain relevant tests for model misspecification. It is essential to ensure that the restrictions imposed on the model are valid. Note that discussion of diagnostic tests concern the EC models' residuals,[1] which is not to be confused with the residuals from the static equations that form the error correction term in the EC models.

The most common and severe problem in time series models is autocorrelated residuals (Wooldridge, 2015). One can assume that the residuals follow the following process:

$$u_t = \rho u_{t-1} + v_t \tag{3.17}$$

If $\rho \neq 0$, the residuals exhibit autocorrelation. Autocorrelated residuals invalidates inference tests in an OLS model. A test to detect autocorrelation is done by testing the residuals. The simplest test is the Durbin-Watson test (DW), which tests the relationship between the residuals and the value of the previous period, also called first-order serial correlation. As such, the Durbin-Watson test is inappropriate when lagged values of the dependent variable are included in the EC model (Brooks, 2008).

The Breusch-Godfrey test is better suited in this case (Wooldridge, 2015). This test is a more general test to uncover any order of serial correlation. Here, the null hypothesis is that the current residual does not have a strong correlation with any of the previous values. If the test statistic exceeds the critical value from the chi-square distribution, the null hypothesis of no serial correlation will be rejected. The Breusch-Godfrey test poses a challenge when it comes to selecting the optimal number of backdated residual values to include in the test. To address this, it's common to base the decision on the frequency of the data.

[1] i.e. the unexplained variation in the model's dependent variable, $(\frac{SST - SSE}{SST})$

For instance, when working with quarterly data, Brooks (2008), advises the use of 4 lags.

It is also essential to test for heteroskedasticity, which implies testing for constant variance in the error term. This can be done using the chi-square-distributed Breusch-Pagan test. If the null hypothesis of constant variance is rejected, it indicates that the error term's variance changes with different values of the explanatory variables (Wooldridge, 2015). With heteroskedasticity, the usual inference tests become invalid, and one may erroneously conclude significance where there is none.

The assumption in the empirical specification of the model is linearity in its estimated parameters. Linearity can be tested with a RESET test (Brooks, 2008). This is a general test for misspecification of the model's functional form. The method uses an auxiliary regression in which the endogenous variable is a function of its predicted values in the n-th order, in addition to other explanatory variables. Predicted values of higher orders can capture variation in the model's nonlinearity. The test for misspecification is performed by estimating an auxiliary regression and testing the null hypothesis with an F- or LM-test. Rejection of the null hypothesis indicates some functional from misspecification.

To check if the model has normally distributed error terms, one can use the Jarque-Bera test. Testing for normality is important for validating both the distribution properties of the OLS estimator and the validity of t- and F-tests. The normal distribution is symmetric around its own mean, and this is precisely what the test examines (Brooks, 2008).

In addition to these tests, it is important to test for possible ARCH effects in the residuals, which refers to autoregressive conditional heteroskedasticity. The presence of ARCH effects indicates that the error term's variance changes over time, which may cause standard errors to be misestimated and lead to incorrect inference. The ARCH test is performed using the LM test, and it detects the presence of ARCH effects in the residuals of the model. It can be thought of as a test for autocorrelation in the squared errors (Brooks, 2008).

3.5.1 Parameter Stability

The parameters we estimate in our models are implicitly assumed to be constant for the entirety of the period in question. This assumption can be formally tested by a Chow test.

The Chow test involves estimating 3 regressions of the same form. One for the whole sample period, and one for each of the two sub-samples. In effect, the sample must be split. Brooks (2008), suggests looking for an obvious structural change in the plot of the time series, or consider any known historical events that could have brought about structural changes for the selection of the split.

The Chow test is appropriate when $T >> k$, where k denotes the number of regressors (Brooks, 2008). This must hold for each of the tree regressions. The null hypothesis is that the estimated parameters are stable across the whole sample. The test statistic is given by:

$$F - stat = \frac{RSS - (RSS_1 + RSS_2)}{RSS_1 + RSS_2} * \frac{T - 2k}{k} \tag{3.18}$$

RSS, RSS_1 and RSS_2 denotes the sum of squared residuals for the whole sample, and the two sub-samples, respectively. The null is rejected if the $F-stat$ is larger than the critical value from the F-distribution with $F(k, T-2k)$. If the null is rejected, separate estimations of the sub-samples should be conducted, or the model should be respecified.

Every aforementioned test will be performed when the error correction models are estimated in Chapter 6.4. For now, we consider the literature on the subject of our analysis and present our specifications.

4 Empirical approaches

This chapter reviews the related literature on the subject of our analysis. After, we provide details on our method and model specification.

4.1 Empirical approaches and related literature

A considerable body of empirical research and modeling approaches has examined the time-varying correlation of returns between equities and treasuries, along with the related effects of inflation and growth. The authors vary among academics and business professionals, and the two categories often intersect. This literature encompasses studies that investigate realized returns as well as those that consider the intrinsic relationship between equities and treasuries, commonly in the form of the Fed model. Different time periods have revealed distinct characteristics due to structural changes.

For the purpose of studying the correlation of returns, one could study the realized returns themselves. For instance, Brixton et al. (2023) recently proposed a model that treats the correlation coefficient of returns between the two asset classes as the dependent variable. According to their findings, rising growth volatility and growth-inflation covariance generate negative correlation between returns. Conversely, inflation volatility is predicted to produce positive correlation, with a considerably stronger magnitude. While their model demonstrates an excellent fit and practical utility, it should be noted that the correlation coefficient is bound between -1 and 1 in general, potentially violating the assumptions of linearity and homoscedasticity in ordinary least squares (OLS) analysis. Consequently, their modeling approach is not viable for our analysis (although one could circumvent this issue by rescaling the data and applying an appropriate function).

Li (2002) developed a similar model that avoids the aforementioned OLS violations by directly modeling the conditional covariance, which is not bound by the [-1, 1] range. Li's research similarly highlights the significant influence of uncertainty surrounding long-term expected inflation on stock-bond correlations, with less impact observed for other macroeconomic factors. Baele et al. (2009) however, finds that macroeconomic factors have limited explanatory power compared to liquidity proxies, albeit within a different modeling framework.

Despite receiving criticism, the Fed model has garnered support for the existence of a cointegrating relationship between treasury yields and earnings yields. Zakamulin and Hunnes (2019) found that the ratio of the treasury yield to the LTM earnings yield was unity between 1958 and 2017, which implies no risk premium. Their evidence suggests that equities follow treasuries in both the long and short run, but not vice versa. Their vector error correction models (VECM) treats every variable endogenously, so the relationship is estimated both ways. They found that a new equilibrium between stock and bond yields emerged in 1958 due to a major paradigm shift in stock valuation theory. Similarly, Koivu et al. (2005) employed a VECM framework and discovered that the Fed model has predictive power, as deviations from equilibrium in the yield ratio indicate corrections back to the long-run relationship between the two yields.

In a representative agent pricing model, Hasseltoft (2009) proposed a solution to puzzles that were previously attributed to irrational behavior, such as inflation illusion. He argued that inflation shocks signal lower future consumption growth, which the agent dislikes. Consequently, the agent demands a positive risk premium for both assets, leading to the co-movement of their yields. Notably, Asness (2002), although famously critical of the Fed model's predictive power, acknowledged its usefulness as a descriptive tool and acknowledged that including a relative volatility measure of the two asset classes in his analysis explains why equities underperformed treasuries from 1950 to 2000. By incorporating this measure, he also observed that the yield ratio approximately equals unity, even going back to the 1920's. Micaletti (2020) argues that although risk premia and volatility adjustments improve the Fed model's performance, they are somewhat arbitrary fixes to a fundamentally flawed model.

Like others have found, Johnson et al. (2013) also finds that both equities and treasuries share a negative sensitivity to inflation shocks. With respect to growth, equities have a positive sensitivity and treasuries has a negative sensitivity. Their Error Correction models include unemployment as a state variable, and the deviation in the Federal Funds rate from the Taylor Rule too. Yet they do not account for potential effects of the inherent relationship between equities and treasuries, even though they also estimate the relationships both ways.

Also, the validity of their analysis is not transparent, as public investment research is not subject to rigorous academic standards.

4.2 Hypothesis, method and empirical specification

As we've seen, there are several ways to evaluate the stock-bond relationship, the correlation of their returns and effects of inflation and growth. In our approach, we evaluate the time series properties of these variables. The long term cointegrating relationships, and the short term dynamics in the ECMs are especially important.

Equation (2.12) showed how the two yields can equal each other under the unreasonable assumptions needed for the Fed Model. As previously discussed, the earnings yield has a tendency to mean-revert, although it was never intended to. The order of integration will be formally tested for both the earnings yield and the treasury yield.

Given that they are I(1), our hypothesis is that there exists a long term equilibrium relationship between the earnings yield and the treasury yield:

$$\frac{\bar{E}_t}{P_t} = \alpha_{11} + \alpha_{12} Y_t \tag{4.1}$$

$$Y_t = \alpha_{21} + \alpha_{22} \frac{\bar{E}_t}{P_t} \tag{4.2}$$

Like in a VECM framework, we estimate this relationship both ways to uncover the causality. Still, (4.1) and (4.2) would capture the same, parsimonious equilibrium. If the earnings yield increases, does the treasury yield follow, and vice versa? Can the causality be reconciled with "The Substitution channel" and "The PV channel"?

By the logic of the Fed Model, α_{12} and α_{22} should equal unity. The two yields should be homogeneous of degree 1, which implies the absence of a risk premium.

The hypothesis of $RP_t = 0$ can be tested by $\alpha_{12} = \alpha_{22} = 1$. Given a rejection of this hypothesis, but cointegration in (4.1) and (4.2), that would imply the existence of a risk adjusted equilibrium. Thus, we formulate an extension. We postulate that the equilibrium is determined by the state variables economic activity and the inflation rate too, through expectations for earnings claims and risk.

If we find that inflation and the unemployment rate follow I(1) processes, we continue with said extension of the equilibrium. In this case, the assumption of $RP_t = 0$ is relaxed. RP_t as an explanatory variable would be interesting. However, the risk premium is unobservable:

1. Solid data is unavailable, other than proxies like INF_t and UR_t.
2. RP_t as an explanatory variable could introduce serious endogeneity and co-linearity issues, because it is conventionally calculated from the expected returns of both the assets in question.

Thus, we formulate the risk-adjusted equilibrium:

$$\frac{\bar{E}_t}{P_t} = \alpha_{31} + \alpha_{32}Y_t + \alpha_{33}inf_t + \alpha_{34}g_t \tag{4.3}$$

$$Y_t = \alpha_{41} + \alpha_{42}\frac{\bar{E}_t}{P_t} + \alpha_{43}inf_t + \alpha_{44}g_t \tag{4.4}$$

Equations (4.3) and (4.4) capture the risk adjusted equilibrium. Given that these 2 cointegrating relationships exist, we would have support for a long term equilibirum in the two assets on a risk adjusted basis. Any deviation from the 2 equilibria will be corrected by changes in the prices of equities or treasuries.

Given cointegration in equations (4.1) to (4.4), we formulate Error Correction models for their first differences that include the respective deviations from equilibrium. Contemporaneous and lagged effects of the explanatory variables can capture short term dynamics and increase the models' explanatory power too. The EC models can be expressed as:

$$\Delta(\bar{E}_t/P_t) = a_1 - a_{E/P}\hat{z}_{t-1} + a_{11}\Delta(\bar{E}_{t-1}/P_{t-1}) + \sum_{i=0}^{I} a_{12}\Delta Y_{t-i} \tag{4.5}$$

$$\Delta(Y_t) = a_2 - a_Y\hat{z}_{t-1} + a_{21}\Delta(Y_{t-1}) + \sum_{i=0}^{I} a_{22}\Delta(\bar{E}_{t-i}/P_{t-i}) \tag{4.6}$$

$$\Delta(\bar{E}_t/P_t) = a_3 - a_{E/P}\hat{z}_{t-1} + a_{31}\Delta(\bar{E}_{t-1}/P_{t-1}) + \sum_{i=0}^{I} a_{32}\Delta Y_{t-i} + \sum_{i=0}^{I} a_{33}\Delta inf_{t-i} + \sum_{i=0}^{I} a_{34}\Delta g_{t-i} + \epsilon_t \tag{4.7}$$

$$\Delta(Y_t) = a_4 - a_Y\hat{z}_{t-1} + a_{41}\Delta(Y_{t-1}) + \sum_{i=0}^{I} a_{42}\Delta(\bar{E}_{t-i}/P_{t-i}) + \sum_{i=0}^{I} a_{43}\Delta inf_{t-i} + \sum_{i=0}^{I} a_{44}\Delta g_{t-i} + \epsilon_t \tag{4.8}$$

ϵ_t is assumed to be white noise, and $\hat{z}_{t-1}$ denotes the deviations from the respective equilibrium. In any deviation from the equilibrium, i.e. $\hat{z}_{t-1} \neq 0$, the error term will affect the growth in the respective yield so that the equilibrium is restored. Such a disequilibrium represents an opportunity for superior risk adjusted returns, market participants will exploit the opportunity and market forces correct the error. a_{E_P} and a_Y capture the speed of adjustment.

Together, the results from the long term equations (4.1 to 4.4) and the ECMs (4.5 to 4.8) will provide insight into the two yields' overall sensitivity to inflation and economic activity. We expect their sensitivities to be aligned with the theoretical background in Chapter 2. Specifically, we expect positive correlation of returns when inflation shocks dominate.

In these specifications, we assume a perfect and frictionless world, meaning that the adjustment process to equilibrium is linear and symmetric, regardless of a positive or negative deviation. Likewise, the effects of exogenous variables are assumed to be linear and symmetric too.

5 Data

The first section of this chapter provides some detail on the data set and the series. The second section presents descriptive statistics, a visual inspection and an informal examination of stationarity properties.

5.1 Dataset

The two endogenous variables in our analysis are:

- $(\bar{E}_t/P_t)$: Inverted Case Shiller Price Earnings Ratio (SP500 index)(Inflation adjusted)
- (Y_t): US10 Year Treasury Yield, constant maturity

Our state variables are:

- (INF_t): Consumer Price Index, quarter over quarter change
- (GDP_t): Unemployment Rate

We use end of quarter (EOQ) observations for all variables. Every series in our data set stretches from 1959Q1 to 2018Q4. The analysis is based on 240 quarterly observations.

We consider this to be the best choice for our time period because we cover a long historical segment that captures the structural inflation of the 70's, and the lack thereof after 1990. This period avoids the breakpoint in 1958 in the stock-bond relationship discovered by Zakamulin and Hunnes (2019), as this could have led to non-stationary series as discussed in Chapter 3. A longer time period would have to account for this structural change somehow.

For shorter and more recent periods, data accessibility might be better in some respects. Forecasts of growth and inflation might capture the dynamics better because markets are forward looking. Also, proxies of a risk premium might be available. For example, the VIX is only available for the last few decades. Even though data is also available from 2018Q4 onwards, it could be subject to more noise. We are able to capture the covid-shock and resurfacing inflation, but this could be a weakness in the analysis, as these shocks have yet to be fully corrected, leading to biased estimators in an Error Correction model. Thus we opt to not include the most recent data in our analysis except for specific graphical illustrations in this chapter.

Logarithmic specifications have been considered, but we find that the variables are best considered in their unchanged state. Firstly, the inflation series contain negative observations, so a logarithmic transformation is not possible there. Moreover, all variables are percentage points, so their unchanged state provides convenient interpretations of results. In this case, we are effectively estimating changes in the dependent variable in terms of percentage points for a one percentage point change in the independent variable. This is akin to estimating elasticities in a log-log specification, even though we are using a lin-lin model.

The quarterly frequency is chosen mostly because of availability issues. Monthly data would be preferable due to more variation captured, but that is unavailable for most of our series going back to 1959. Still, a quarterly frequency should be enough to capture short term variation in our series. The quarterly frequency might be the highest frequency at which our models would have explanatory power. The quarterly frequency is commonly chosen in similar analyses[1].

5.2 Descriptive statistics

In this section, we consider the evolution of our time series and try to discern their stationarity properties. First we consider the two yields and their rolling correlation. After, we consider the inflation and unemployment series, and finally the autocorrelation functions for every series. We report a correlation matrix in A.2 and descriptive statistics in A.1.

5.2.1 Equity and Treasury Yield

Figure 5.1 plots the two yields in levels along the left axis, and their 5-year rolling correlation on the right axis.

We note that the correlation is mostly in positive territory throughout the 70's and 90's. The 2000's exhibited negative correlation, apart from a brief period around 2015 and recently again in 2022.

[1]For example, Baele et al. (2009), Johnson et al. (2013), and Zakamulin and Hunnes (2019) use quarterly data.

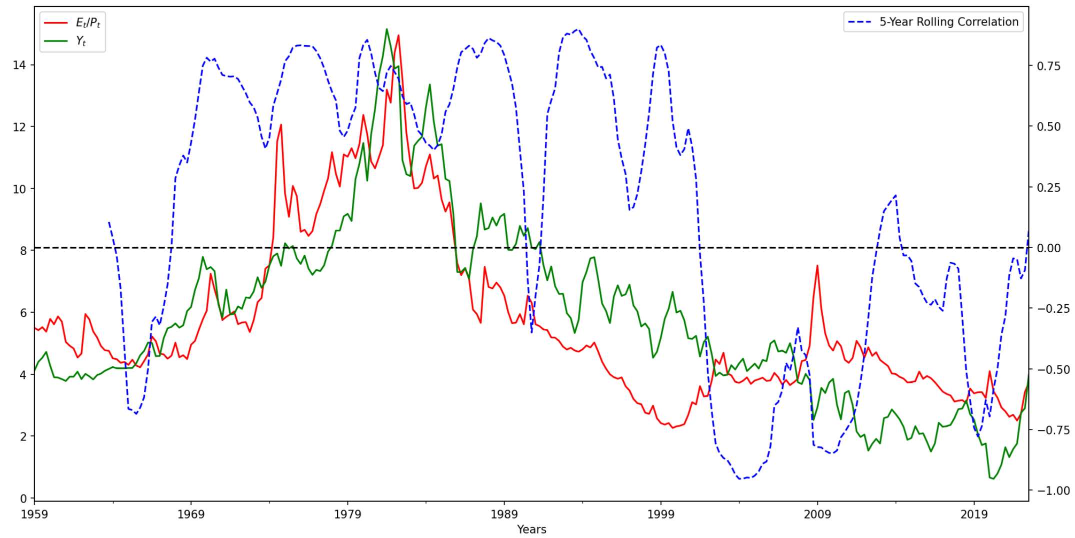

FIGURE 5.1: $\bar{E}_t/P_t$ and Y_t.

The two yields are for the most part comoving, except during the dot-com bubble until the GFC where the earnings yield trends against the treasury yield. There are large deviations around the 2001 dot-com bubble, the sovereign debt crisis in 2011 and the covid-19 shock. From 1959 they seem to comove closely and interchangeably in an upward trend. The trend reverses around 1984, after which the treasury yield rests above the equity yield. This dynamic flips around the mid 2000's, and the equity yield rests above the treasury yield. The persistence of the deviations, and the flipping dynamic allude to other variables exerting an effect.

The yields' first differences are of interest, and their evolution is plotted in Figure 5.2. Time series with stochastic trends, like Y_t and $\bar{E}_t/P_t$, contain at least 1 unit root. If the two yields follow I(1) processes, their first differences should not exhibit any trend following or random walk.

As explained in Chapter 2, the first differences of the yields are approximations of returns. It is not surprising then, that the first differences seem to exhibit some time-varying volatility. Still, they vary evenly on either side of the zero-line. They appear to mean-revert. The first differences correspond nicely to our suspicion that (Y_t) and $(\bar{E}_t/P_t)$ follow I(1) processes.

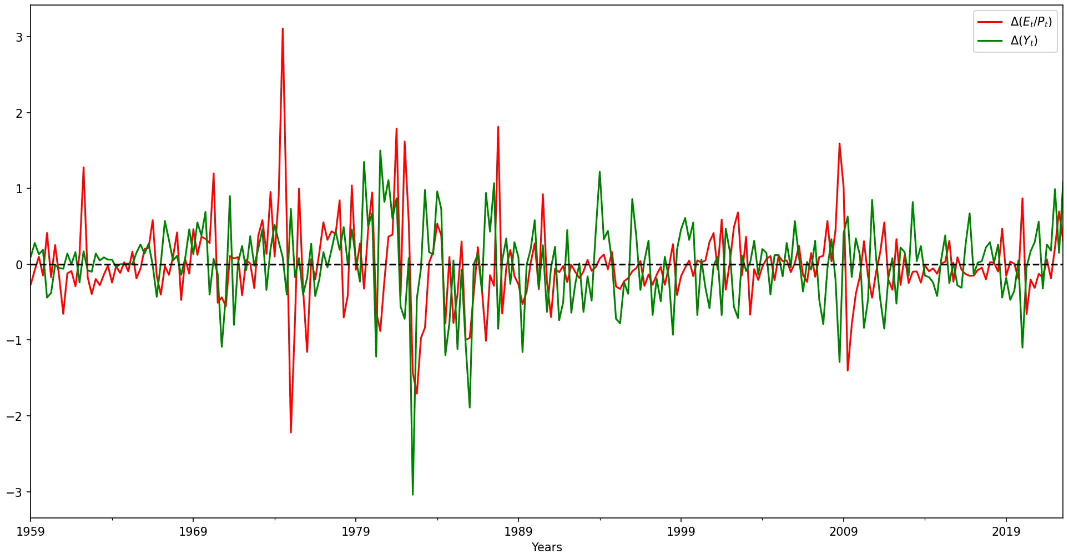

FIGURE 5.2: $\Delta(\bar{E}_t/P_t)$ and $\Delta(Y_t)$.
The rolling correlation of returns is omitted because the variation becomes hard to make out, but the pattern remains.

5.2.2 Inflation and Unemployment

The stationarity properties of our state variables must be determined too. Upon a visual inspection, our inflation and unemployment series in Figure 5.3 seem to somewhat co-trend with the two yields. All 4 series seem to trend upwards until the early 1980's, and then reverse their trends downward. The GFC produced the only instance of negative values in the inflation series.

The inflation and unemployment series exhibit more mean reversion properties than the two yields. UR_t especially is characterized by sudden shocks that dissipate with time. As for inflation, it is the other way around; it tends to trend upwards and then abruptly drop. Because of the trends and mean reversion, it is not obvious if the inflation and activity series follows I(1) or I(0) processes.

Their first differences in Figure 5.4 also vary evenly around zero, and indicate I(1) processes like the two yields.

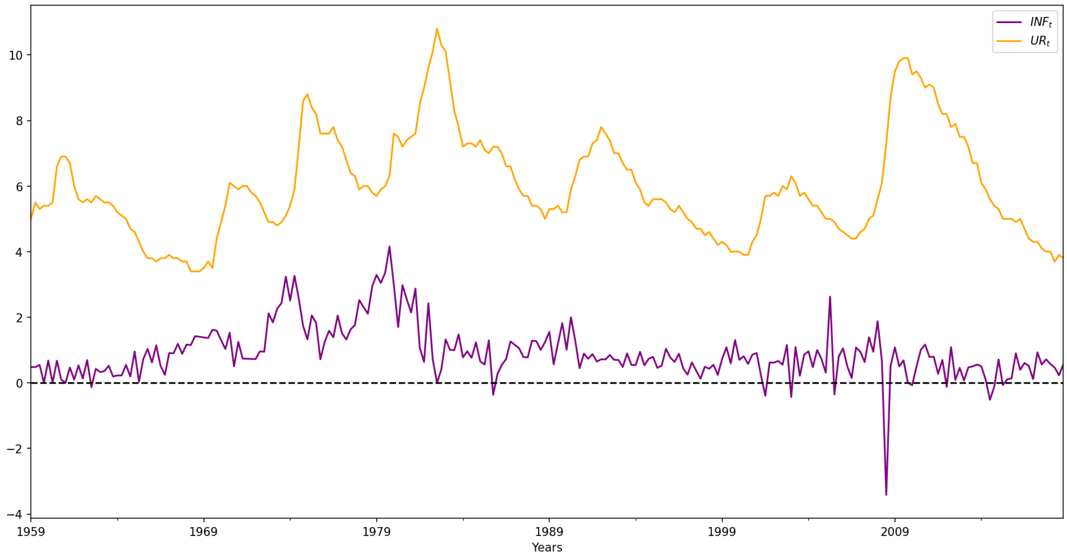

FIGURE 5.3: INF_t and UR_t. Data from 2019 Q4 onwards is omitted from plot because extreme shocks during covid make it difficult to make out the variation throughout the period.

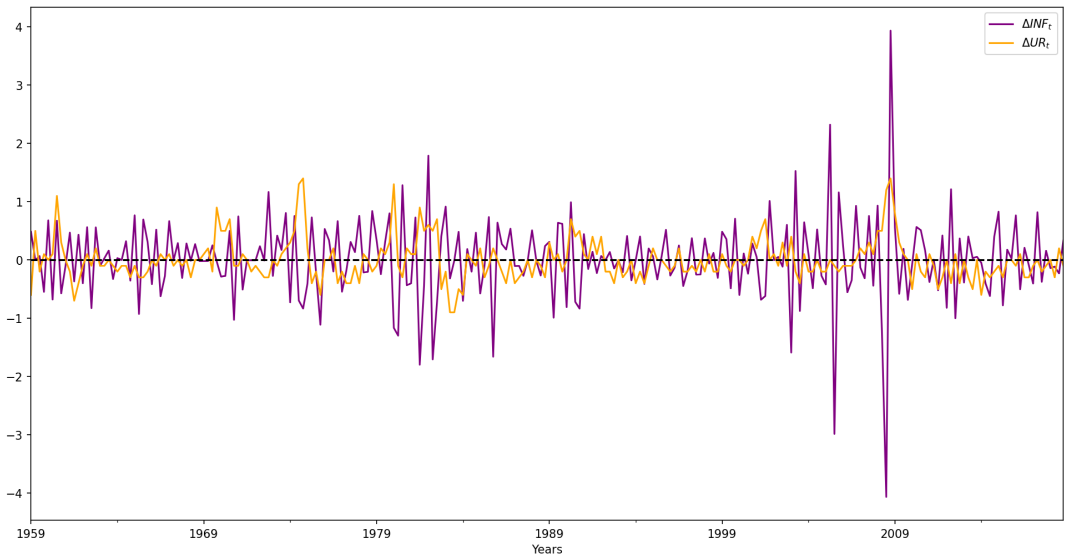

FIGURE 5.4: ΔINF_t, ΔUR_t. Covid data is omitted here too.

6 Empirical analysis

In accordance with Chapter 3, this analysis will consider the time series properties of our series formally to begin with. We estimate long run steady state equations and test for cointegration. Next, we estimate short and long run relationships using 2-step Error Correction models, and continue said testing. Finally, we discuss the estimates and their implications. Put simply, the objective is to study the correlation of returns between equities and treasuries in different economic conditions, while controlling for an inherent causal relationship between them.

6.1 Testing for stationarity

As mentioned earlier, all 4 series must be integrated of the same order for an EC specification to be valid. We formally test each series for weak stationarity in this section. We conduct somewhat exhaustive ADF-tests as described in Chapter 3. We use equation (3.10) and substitute y_t for our 4 variables of interest.

The selection of lag length q in ADF-tests can prove challenging. Excessive lags reduce the power of the test (Enders, 2014). A trade-off exists between lags that correct for serial correlation and fewer degrees of freedom that lead to biased estimators. Several approaches are available in lag selection. One is to use an information criterion that considers this trade-off formally, e.g. equation (3.12). Another is to consider the statistical significance of the parameters in front of the lagged first differences. Enders (2014) describes a *general to specific* methodology which we employ.

1. We estimate (3.10) for each variable with a long lag length q^*, in our case 15.

2. We check the t-statistic of the last lag q^*. If it is insignificant at the relevant critical value, we re-estimate the regression with a lag length $(q^* - 1)$.

3. We repeat the process until the last lag is significantly non-zero.

This iterative procedure leaves us with the optimal lag length to use in the subsequent ADF-tests. We go through the procedure for each variable and report the optimal lag length q in Table 6.1 below.

We formally test the variables for their stationarity in levels, and include their respective lagged first differences q. A constant term is also included in the test. The null hypothesis of the test is non-stationarity, against the alternative of weak stationarity.

Standard t-values are not valid under the null, so other critical values are used. The test is conducted through Python Statsmodels' adfuller package, which "... [uses] surface approximation from MacKinnon 1994, but using the updated 2010 tables" (Statsmodels, n.d) in the selection of critical values. As seen in Table 6.1, the null cannot be rejected for any variable.

Variable	**Lag (q)**	**t-value**	**p-value**	**Conclusion**
$(\bar{E}_t/P_t)$	2	-1.46	0.566	Non-stationary
(Y_t)	13	-1.19	0.677	Non-stationary
(INF_t)	15	-1.90	0.323	Non-stationary
(UR_t)	8	-2.59	0.096	Non-stationary

TABLE 6.1: ADF tests of variables.

In order to determine the variables' order of integration, we conduct the same procedure on the first differences of the variables. As mentioned, a cointegrating relationship between the variables require them to be integrated of the same order. The nullhypothesis is, again, non-stationarity against the alternative of weak stationarity. This test also includes a constant term. The results are reported in Table 6.2.

The tests reject the null for every first difference. All first differences are stationary, suggesting every variable is I(1). They all meet this requirement for cointegration.

Variable	Lag (q)	t-value	p-value	Conclusion
$\Delta(\bar{E}_t/P_t)$	15	-3.20	0.020	Stationary
$\Delta(Y_t)$	12	-4.03	0.001	Stationary
$\Delta(INF_t)$	12	-6.52	0.000	Stationary
$\Delta(UR_t)$	7	-6.08	0.000	Stationary

Table 6.2: ADF tests of first differences.

6.2 Cointegration analysis

In accordance with chapters 3 and 4, we expect a long term stationary relationship between the two yields, inflation and unemployment rate. All 4 series are integrated of the first order. Now we investigate if said long term relationships exhibits stationary residuals in the second step. We investigate these potential cointegrating relationships in the time period 1959Q1-2018Q4.

We consider combinations with $(\bar{E}_t/P_t)$ and (Y_t) as the endogenous variables. In a VECM framework we would consider every possible combination of cointegrating relationships, but not here. Equations with (INF_t) or (UR_t) as endogenous variables and the yields defined exogenously don't make much sense in any context.

In order to test for cointegration, we first need to evaluate the steady state, long run equations, (4.1) to (4.4). This is done by a standard OLS procedure using the python Statsmodels package. The regression results are reported as (6.1) to (6.4) in Table 6.3 below. These equations meet the first critera for cointegration; they are all I(1). The second criteria, for stationary residuals, will be tested in the second step. In effect we are performing a 2-step Engle Granger test.

In the second step, we test whether the residuals from the static equations are stationary by ADF tests. If they are, the relevant equation meets both criteria for cointegration. We substitute the residuals for $\hat{u}_t$ in the equation below.

$$\Delta\hat{u}_t = \theta\hat{u}_{t-1} + \sum_{i=1}^{q}\pi_i\Delta\hat{u}_{t-i} + v_t$$

The nullhypothesis is that the residuals are non-stationary. In effect: $H_0 : \theta = 0$ against $H_1 : \theta < 0$.

TABLE 6.3: Cointegrating equations. 1959Q1-2018Q4

	$(\bar{E}_t/P_t)$		(Y_t)	
Explanatory Variables	(6.1)	(6.3)	(6.2)	(6.4)
$(\bar{E}_t/P_t)$			0.8401	0.9470
			(0.000)	(0.000)
(Y_t)	0.7449	0.5134		
	(0.000)	(0.000)		
(INF_t)		0.7696		0.1538
		(0.000)		(0.364)
(UR_t)		0.6862		−0.3752
		(0.000)		(0.000)
$Const$	1.3361	−2.0670	1.1526	2.6320
	(0.000)	(0.000)	(0.000)	(0.001)
Observations	240	240	240	240
R-squared	0.626	0.793	0.626	0.661
F-statistics	398.1	301.0	398.1	153.5

Note: p-values in parentheses ()

For convenience, we rely on information criteria in the selection of q instead of using the previously employed general to specific methodology. Specifically we are using **Akaice's Information Criterion**, equation (3.12). Neither a constant or a trend is used in the test, because the residuals are centered around 0. The results are reported in Table 6.4.

Equation	Lag	t-value	Critical value	Conclusion
(6.1)	2	-2.90	-1.94	Stationary
(6.2)	2	-2.85	-1.94	Stationary
(6.3)	7	-2.26	-1.94	Stationary
(6.4)	2	-2.70	-1.94	Stationary

TABLE 6.4: Engle-Granger

The Engle Granger tests supports stationary residuals for every static equation. The nullhypotheses of non-stationary residuals are discarded at 5% significance level. Thus both criteria for cointegration are met in (6.1) to (6.4).

By the Granger representation theorem, cointegrating relationships exhibit error correction. Thus we model our 2-step EC-models in the next section after discussing the long term relationships.

6.3 Long term relationships

First, we briefly discuss the results in terms of the restrictive Fed Model. Next, we consider the long term dynamics of the equilibria, first by discussing the inherent stock-bond relationship, and then the effects of macroeconomic state variables.

6.3.1 Fed Model

In the context of the Fed Model, we consider the estimates in (6.1) and (6.2), shown in Table 6.3.

By the Fed Model, $\alpha_{12} = \alpha_{22} = 1$ if the risk premium equals 0 in the long term. In that case the two yields are homogenous of degree 1. That means a 1 percentage point increase in either yield causes a proportional change in the other one. If $RP_t > 0$, $\alpha_{12} > 1$ and $\alpha_{22} < 1$.

We are unable to test the hypothesis jointly, so separate t-tests are conducted. Equations. 6.1 and 6.2 estimates both parameters to be below unity. As such, we obtain conflicting results. We do not find support for the Fed Model's restrictive assumption of $E[RP_t] = 0$, like Zakamulin and Hunnes (2019), did, but we cannot reject it either, by (6.1) and (6.2). By equations (6.3) and (6.4), we cannot test it explicitly. The failure to reject the hypothesis could be due to the earnings yield being biased downward, as we noted in Chapter 2.

However, the cointegration in (6.1) through (6.4) is evidence of long term equilibria, which is not restricted by $RP_t = 0$. We are primarily interested in the risk-adjusted equilibrium given by (6.3) and (6.4).

6.3.2 Long term relationship - risk adjusted

In (6.3) and (6.4), the state variables capture the risk adjustment, i.e. the unobservable risk premium through the state variables INF_t and UR_t. Thus the economic interpretation of α_{12} and α_{22} are not the same as for α_{32} and α_{42}.

Notably, the long term causality across the two yields is bidirectional in (6.3) and (6.4). This means equities and treasuries follow each other in the long term. When the yields of the two assets deviate too far from one another, market forces capitalize on the opportunity and restore the equilibrium. This supports that "The Substitution channel" works in the long term.

Moving on, we discuss the effects of the state variables. Note that the explanatory power when controlling for inflation and economic activity rises meaningfully in (6.4), but even more so (6.4). Both regressions' F-statistics fall, yet they are still extremely high. In terms of magnitude, the effect of the Y_t falls and $\bar{E}_t/P_t$ rises when the equilibrium is extended.

Equation (6.3) indicates the earnings yield has a strong sensitivity to the inflation level, both in terms of statistical significance and magnitude. In real terms, equity prices suffer when inflation increases. For some intuition on the magnitude, a 1 percentage point jump in the long term inflation rate will increase the long term expected earnings yield by 76.96 basis points[1].

Equation (6.3) indicates the earnings yield is sensitive to the unemployment rate too, as expected. This means that when the long term activity level falls, earnings are likely to suffer in real terms, which the market prices in.

The treasury yield exhibits no sensitivity to the inflation level, surprisingly, in (6.4). The treasury yield rises when the activity level falls in the long term. In that case, the risk free asset becomes more favorable than the risky one.

To summarize, this subchapter found support for a risk-adjusted equilibrium between the two assets. Thus, we moved on with our analysis in the context of a risk adjusted equilibrium. We discussed the long term effects in (6.1) to (6.4). Next, we estimate their respective EC model specifications that capture the error correction mechanism and short run effects.

[1] $\frac{E_t}{P_t}$ of 1 translates to 100 basis points. Although this sensitivity can seem excessive, keep in mind that 4 successive quarters of 1% inflation translates to 1.01^4 => 4.06% annualized inflation, twice that of the inflation target.

6.4 Error Correction models

TABLE 6.5: Error Correction models. 1959Q2-2018Q4

	$\Delta(\bar{E}_t/P_t)$		$\Delta(Y_t)$	
Explanatory Variables	Model 1	Model 3	Model 2	Model 4
$\Delta(\bar{E}_t/P_t)$			0.0785	0.0443
			(0.230)	(0.504)
$\Delta(\bar{E}_{t-1}/P_{t-1})$	0.1879	0.1154	−0.0739	−0.0038
	(0.003)	(0.079)	(0.249)	(0.955)
$\Delta(Y_t)$	0.0772	0.0443		
	(0.232)	(0.490)		
$\Delta(Y_{t-1})$	0.1629	0.1669	0.1136	0.1190
	(0.013)	(0.008)	(0.088)	(0.064)
$\Delta(INF_t)$		0.1703		0.0835
		(0.001)		(0.107)
$\Delta(INF_{t-1})$		0.1444		0.2359
		(0.006)		(0.000)
$\Delta(UR_t)$		0.6147		−0.1168
		(0.000)		(0.306)
$\Delta(UR_{t-1})$		−0.1807		−0.1708
		(0.088)		(0.113)
$\hat{z}_{t-1}$	−0.0693	−0.0833	−0.0435	−0.0561
	(0.004)	(0.002)	(0.029)	(0.005)
Const	−0.0051	−0.0048	−0.0071	−0.0068
	(0.879)	(0.875)	(0.829)	(0.829)
Observations	239	239	239	239
R-squared	0.102	0.248	0.038	0.180
F-statistics	6.671	9.504	2.289	6.310
Breusch-Godfrey	$\chi^2(4)$= 19.71	$\chi^2(4)$= 15.72	$\chi^2(4)$= 5.97	$\chi^2(4)$= 2.16
	(0.001)	(0.003)	(0.201)	(0.706)
Jarque-Bera	$\chi^2(2)$= 459.60	$\chi^2(2)$= 132.16	$\chi^2(2)$= 189.90	$\chi^2(2)$= 76.62
	(<0.000)	(<0.000)	(<0.000)	(<0.000)
Ramsey RESET	F(4,234)=4.43	F(8,230)=8.61	F(4,234)=5.73	F(8,230)=5.07
	(0.109)	(0.013)	(0.057)	(0.079)
Breusch-Pagan	$\chi^2(4)$= 16.58	$\chi^2(8)$= 31.55	$\chi^2(4)$= 13.30	$\chi^2(8)$= 34.76
	(0.002)	(<0.001)	(0.010)	(<0.000)
ARCH	$\chi^2(1)$= 34.33	$\chi^2(1)$= 31.68	$\chi^2(1)$= 35.01	$\chi^2(1)$= 55.19
	(<0.000)	(<0.001)	(<0.000)	(<0.000)
Chow	F(5,229)=1.30	F(9,221)=1.96	F(5,229)=1.17	F(9,221)=2.10
	(0.265)	(0.045)	(0.322)	(0.030)

Note: p-values in parentheses ()
Diagnostic tests are explained in Chapter 3.5.

We present 4 models in Table 6.5, 2 for each yield of the yields. Models 1 and 2 are more parsimonious, capturing the stock-bond dynamics alone. Models 3 and 4 include economic activity and inflation as state variables too. Every model captures deviations from its respective long term equilibrium, denoted $\hat{z}_{t-1}$, which are the residuals from the static equations in Table 6.3. The first lag for every explanatory variable is also added to capture the short term dynamics better. Specifically, the unemployment rate is often considered a lagging indicator of economic activity.

By the Granger representation theorem, cointegrating relationships can be modeled as EC-models, and vice versa. That means our EC-models also serve as an informal Engle Granger test. A statistically significant error correction mechanism indicates cointegration. Note that every ECM is consistent with the EG tests in Chapter 6.2, in that the ECMs exhibit significant error correction terms.

Thus we have consistent results that equilibria given by eqs. (6.1) to (6.4) hold in the long term. First, we discuss results concerning the error correction mechanisms and the inherent stock-bond relationship. After, we consider the effects of the state-variables in terms of correlation regimes in returns.

6.4.1 Error Correction and the stock-bond relationship

The models all exhibit slow adjustment speeds toward their respective equilibria. For example, Model 4 suggests only 5.61% of a deviation from the equilibrium in (6.4) in a given quarter will be corrected by the subsequent quarter. All models indicate a similar magnitude of error correction, although both models for the treasury yield indicate slower adjustment speeds than that of the earnings Yield.

We also note the short term dynamics of the stock bond relationship. Models 1 and 3 suggest equities follow treasuries in the short term, but Models 2 and 4 indicate treasuries do not follow equities. This is supported by the dramatic jump in explanatory power from Model 2 to 4, 3.8% to 18%, meaning the equity yield alone does a poor job of explaining the short term dynamics of the treasury yield. The F-statistic more than doubles from Model 2 to 4.

Why is the causality going one way only in the short term? Unlike the long term effect, this result is consistent with results from Zakamulin and Hunnes (2019). In Chapter 2.4 we established two channels of causality that can provide an explanation.

This one-way causality is consistent with "The PV channel". Higher treasury yields increase borrowing costs for companies, reducing their profitability, and the discount rate increases in a PV context. An efficient market would price this in swiftly, thereby causing negative returns in equities.

6.5 Correlation regimes

In this section we discuss our results in the context of correlation regimes. Do our model estimates coincide with the findings in Johnson et al. (2013)? Specifically, will dominating inflation shocks lead to positive correlation of returns between equities and treasuries? In the absence of such shocks, will the correlation be negative? To answer this, we consider Models 3 and 4 in Table 6.5 that capture short term effects.

In Chapter 2, we showed how the dependent variables are approximations of returns in the two asset classes. $\Delta\bar{E}_t/P_t$ captures the *inverse* of returns in the S&P 500 index. ΔY_t captures the *negative* of returns in the 10 Year U.S. Treasury note, per unit of duration. The approximation is more accurate for $\Delta\bar{E}_t/P_t$. In fact, its correlation with realized quarterly returns in the S&P 500 index is 0.86, as seen in A.1. The actual returns and $\Delta\bar{E}_t/P_t$ are plotted in A.2. Although the correlation is not perfect, it is within the margin of error for the scope of this paper. The imperfection is likely due to the variation that earnings account for.

The correlation regimes in question can be long-lasting, as Figure 5.1 illustrates. Thus we must consider the yields $\bar{E}_t/P_t$ and Y_t themselves. Specifically, we must consider their respective long term cointegrating equations too, as shown in Table 6.3.

Overall, we expect the yields (and returns) of the two assets to share a positive (negative) sensitivity to inflation, and opposite sensitivities to economic activity. In accordance with chapter 2, we expect Y_t to have a positive (negative) sensitivity to the activity level, and $\bar{E}_t/P_t$ a negative (positive) one.

We reiterate the long term sensitivities here. In Chapter 6.3 we found that the earnings yield $\bar{E}_t/P_t$ had a strong positive sensitivity to economic activity, through UR_t, and the inflation rate INF_t. This means a persistent jump in the inflation rate, or a persistent drop in the activity level will, all else equal, induce negative returns in equities through higher $\bar{E}_t/P_t$.

On the other hand, the treasury yield showed, surprisingly, no sensitivity to inflation in the long term. The treasury yield exhibited the expected positive sensitivity to the unemployment rate, meaning a negative sensitivity to economic activity. Thus, a long term jump in the activity level will cause negative returns in treasuries through higher Y_t.

In the short term, the returns of both assets show strong negative sensitivities to inflation growth. Overall, the magnitude is slightly larger for equities than treasuries:

$$\sum_{i=0}^{I} a_{33} = 0.1703 + 0.1444 \quad \text{versus} \quad \sum_{i=0}^{I} a_{43} = 0.2359$$

Only the effect of lagged inflation growth is statistically non-zero for returns in treasuries.

Unlike that of the long term treasury yield, treasury returns are not sensitive to changes in the activity level in the short term. As for equity returns, the contemporaneous effect of growth in the activity level is positive and has, by far, the largest magnitude among every estimate.

The table below summarizes the long and short term sensitivities of macroeconomic effects on returns in the two assets.

Variable	Time-frame	Equities	Treasuries
Inflation rate	Short	-	-
	Long	-	/
Activity level	Short	+	/
	Long	+	-

TABLE 6.6: Return sensitivities
/ denotes no sensitivity, or not applicable
Higher unemployment rate UR_t corresponds to a lower activity level in the economy.

As expected, higher economic activity results in positive returns for equities, but negative for treasuries. Higher inflation pressures induce negative returns for both assets. Thus, these dynamics coincide with those of Johnson et al. (2013) overall. At the very least, our models display no statistically significant effects that contradict those expectations. As such, our models also indicate that when inflation shocks dominate, positive correlation regimes in equities and treasuries will arise. Absent of such shocks, we can expect negative or no correlation between returns in the S&P 500 index and the 10 Year U.S. Treasury note.

6.6 Diagnostic tests

Upon examination of the model misspecification tests, several pertinent observations emerge regarding the properties and issues with Models 1 through 4. Detailed histograms of the models' residuals can be found in A.4.

The Breusch-Godfrey (BG) test indicates autocorrelation in the error terms of Models 1 and 3. This occurrence raises significant concerns regarding the validity of any conclusions drawn from these models. Conversely, the null hypothesis of no autocorrelation is not rejected for Models 2 and 4, suggesting that their residuals do not suffer from this issue. Intriguingly, only the models for equity returns seem to present autocorrelated errors, while those for treasury returns do not.

The Jarque-Bera tests yield results that reject the null hypothesis of normal distribution for all models. It is important to note that such a finding is rather common for return series, the type of data our dependent variables attempt to approximate. Normally distributed returns is an unrealistic assumption (Brooks, 2008). This suggests that the error terms in these models do not adhere to a normal distribution, thereby potentially affecting the validity of t- and F-tests.

The Ramsey RESET tests applied to the parsimonious Models 1 and 2 yield F-statistics and p-values that do not reject the null hypothesis of the correct functional form. This outcome suggests that the linearity assumption might hold for these models. However, the extended models yield contrasting results. Specifically, the null hypothesis of the correct functional form is not rejected for Model 4, while it is rejected for Model 3. This finding suggests that the functional form is valid for treasury returns, but not for equity returns. It appears Model 3 fails to account for some underlying non-linear dynamics specific to equities.

The results of the Breusch-Pagan tests for all models suggest the existence of heteroskedasticity, as indicated by test statistics and p-values that reject the null hypothesis of homoskedasticity. This reveals that the error term's variance changes with different values of the explanatory variables. Even though the Ramsey RESET test does not reject the null hypothesis of correct functional form for Models 1, 2, and 4, the presence of heteroskedasticity, as identified by the Breusch-Pagan test, might be indicative of other model specification issues, such as omitted variables, incorrect functional form, or non-linearity in the relationships between variables.

Finally, the results of the ARCH tests for each model are indicative of volatility clustering in the dependent variables. Given that the dependent variables are approximations of returns, this is not an unexpected finding. ARCH effects are a common feature of return series (Brooks, 2008). The presence of ARCH effects might lead to misestimated standard errors and incorrect inference.

In summation, the results present a mixed picture. Models 1, 2, and 3 seem to satisfy the linearity assumption. For Models 2 and 4, the BG test suggests their residuals are not serially correlated, but the ARCH test suggests the presence of volatility clustering. All four models suffer from issues with non-normality, heteroskedasticity, and ARCH effects. These findings warrant further refinement of the models. An implementation of GARCH-like methodologies within an ECM framework would be interesting. Regardless, the test results reveal a lack of robustness in the models, and issues common to the the study of return series.

6.6.1 Parameter Stability

In line with the explanation provided in Chapter 3.5.1, a Chow test was executed for each model. The test's objective is to examine the stability of the coefficients throughout the entire time period, with the null hypothesis being rejected if the test statistic surpasses the critical value.

The two yields, inflation and the unemployment rate exhibited an upward trend until they reached a peak in the early 1980s. Their first differences exhibit volatility clustering around this time. This peak coincides with an aggressive monetary policy stance taken by Paul Volcker, the then Chairman of the Federal Reserve, to mitigate inflationary pressures in late 1980. The ensuing change in inflation expectations, likely from high to low, serves as the basis for selecting 1981Q1 as the split date for the Chow test.

In order to ensure the validity of the Chow test, it is required that the sample size (T) is substantially larger than the number of parameters (k) in each sub-sample regression. For all models, this requirement is satisfied. The shortest sub-sample has a size of $T = 89$, and for Models 3 and 4, which contain the most parameters, $k = 9$. The findings of these tests are presented in Table 6.5.

Interestingly, the null hypothesis of stable parameter estimates is not dismissed for the less complex Models 1 and 2. However, when short-term dynamics of the state variables are incorporated into Models 3 and 4, the null hypothesis is rejected. As per the recommendation of Brooks (2008), further model adjustments are then in order.

Summary

First, ADF tests indicated that all 4 series were integrated of order 1, so the long term cointegrating relationships were estimated. Engle-Granger tests indicated significant cointegration. The parsimonious relationships produced mixed and unreliable results concerning the risk premium, but we obtained evidence of a risk adjusted equilibrium in the two yields by the more specific specifications determined by economic state variables. Inflation and economic activity did a reasonable job of explaining the long term relationships.

We explained the varying causal relationship between the treasury yield and the earnings yield themselves. Effects of inflation and unemployment, as a proxy for economic activity, had effects on returns and the yields that met our suspicions. We found that changes in economic activity induces a negative correlation of returns in the assets, whereas inflation pressures will induce positive correlation.

7 Conclusion

The primary purpose of this paper was to investigate the different correlation regimes in returns between equities and treasuries caused by fundamental economic state variables.

We have approached this study by using time series methods. We considered the earnings yield of the S&P 500 Index, and the 10-Year U.S. Treasury yield, from which we could derive approximations of their realized returns. We studied the effects of economic state variables, specifically the unemployment rate and the inflation rate, both in the long and short run. Our framework called for a discussion of the contentious Fed Model's implications.

The 4 time series in question were all integrated of order 1 according to our Augmented Dickey Fuller tests. Next we formulated 4 long term relationships, 2 parsimonious ones determined only by the other asset, and 2 more specific determined by the state variables too. This was done in 2's to capture the causal relationship between stocks and bonds. Engle-Granger tests indicated that every relationship cointegrated.

The regressions revealed bidirectional causality between stocks and bonds in the long term, probably due to what we called "The Substitution channel" for convenience. After, we formulated Error Correction models for each cointegrating relationship. The existence of risk-adjusted equilibria was then cemented, and the analysis revealed a unidirectional causality for stocks and bonds in the short term, through the analogous "PV channel".

The effects of the state variables were discussed on a long- and short-term basis, in terms of the correlation regimes in stocks and bonds. The results corroborated earlier findings, in that inflation pressures cause positive correlation of returns between stocks and bonds. Absent such pressures, negative correlation regimes are likely to dominate.

We found that the parsimonious models exhibited stable parameters over the 60-year period in question, unlike the specific models. Robustness tests also indicated, unsurprisingly, that the every model suffered from robustness issues common to return series, such as volatility clustering. A resolution to these issues, that also fits the methods we used, would be interesting.

A Appendix

A.1 Descriptive statistics of time series

Table A.1: Descriptive statistics

Variable	count	mean	std	min	max
$(\bar{E}_t/P_t)$	240.0	5.865711	2.682127	2.262955	14.947683
(Y_t)	240.0	6.080500	2.848316	1.500000	15.150000
(INF_t)	240.0	0.912313	0.813925	-3.416990	4.161250
(UR_t)	240.0	5.987083	1.594433	3.400000	10.800000
$\Delta(\bar{E}_t/P_t)$	240.0	-0.009273	0.538449	-2.220206	3.110308
$\Delta(Y_t)$	240.0	-0.005458	0.524272	-3.040000	1.500000
$\Delta(INF_t)$	240.0	0.002366	0.719631	-4.067220	3.935920
$\Delta(UR_t)$	240.0	-0.007500	0.353976	-0.900000	1.400000

A.2 Correlation matrix

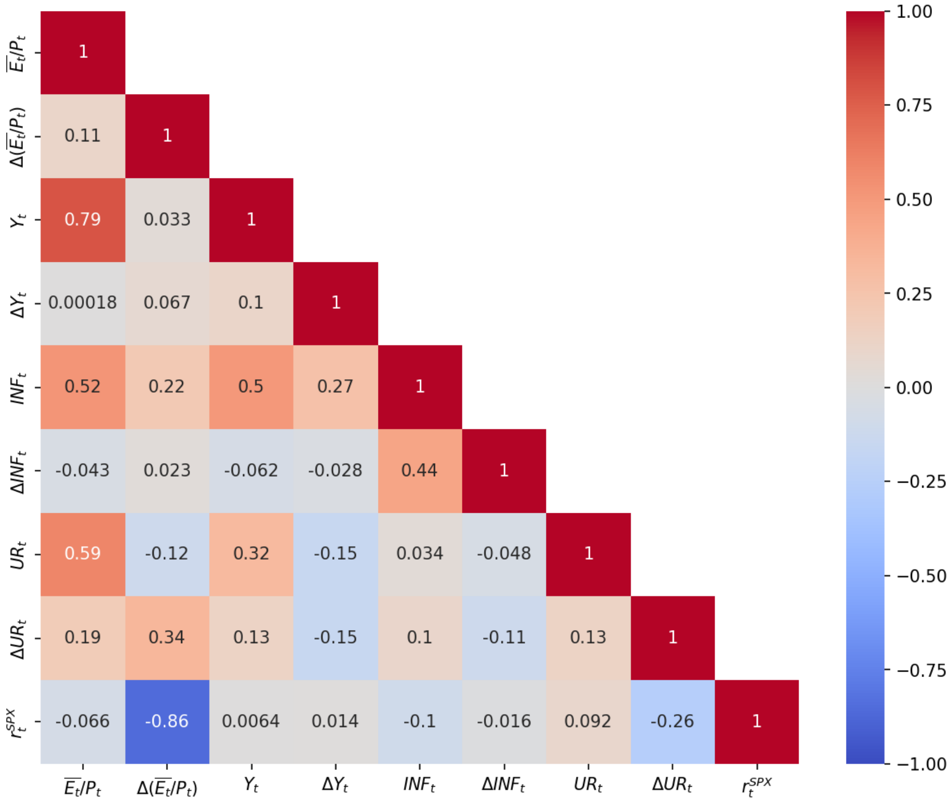

FIGURE A.1: Correlation Matrix and Heatmap. Full period sample: 1959Q1 - 2018Q4
r_t^{SPX} denotes quarterly returns in the S&P500 Index

A.3 S&P 500 quarterly return, and $\Delta\frac{\bar{E}_t}{P_t}$

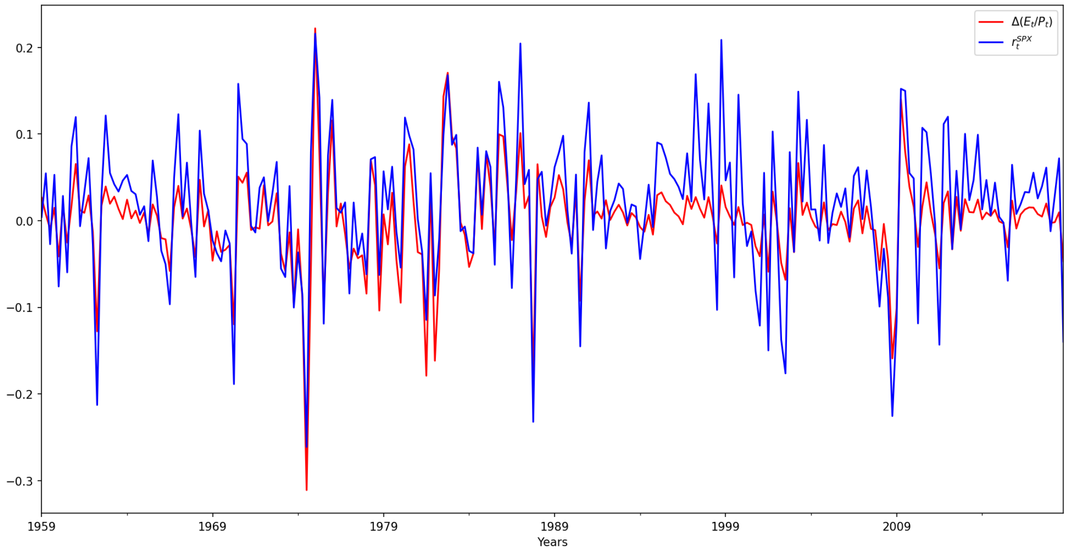

FIGURE A.2: $Corr(r_t^{SPX}, \Delta\frac{\bar{E}_t}{P_t})$

A.4 Model residuals

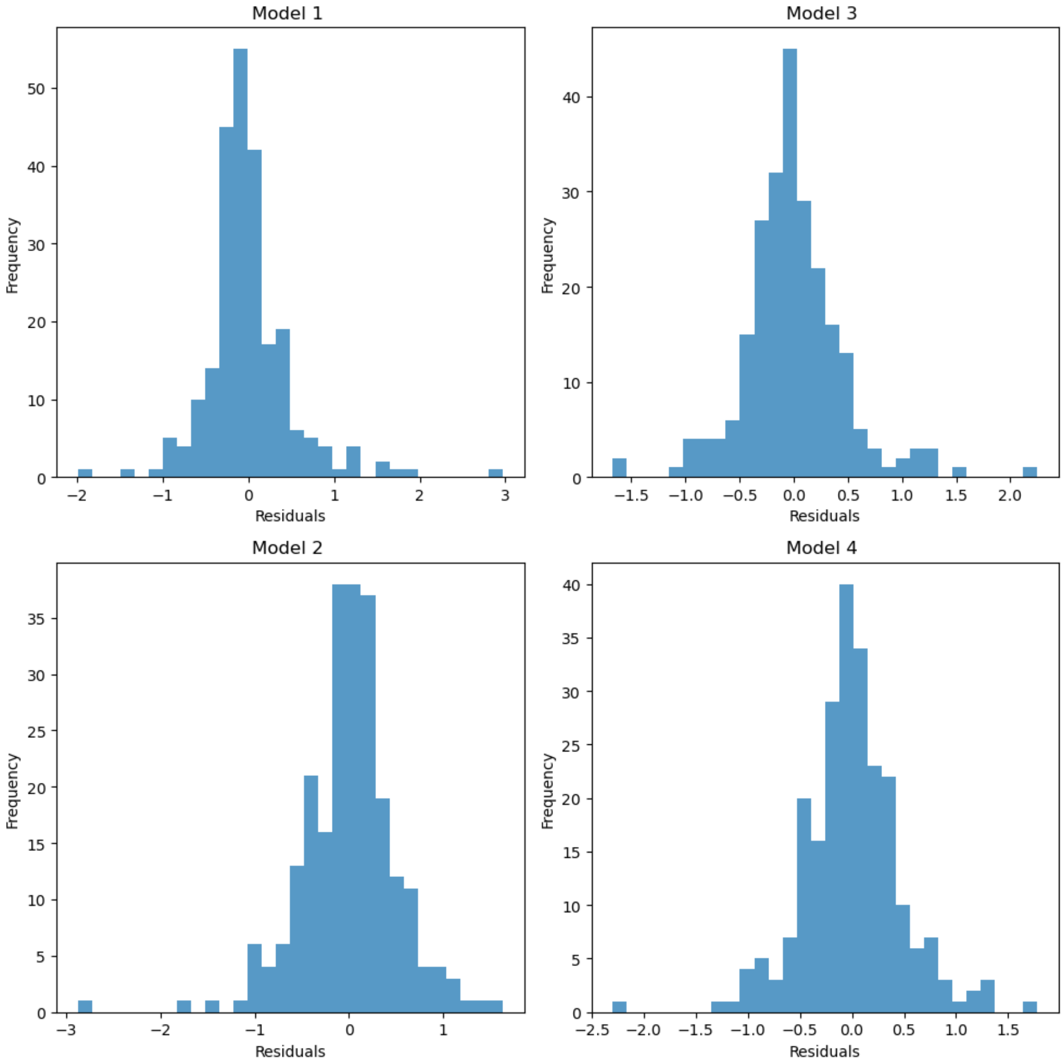

FIGURE A.3: EC model residuals - histogram.
Top row dependent variable: $\Delta\frac{\bar{E}_t}{P_t}$
Bottom row dependent variable: ΔY_t

Bibliography

Angelini, N., Bormetti, G., Marmi, S., & Nardini, F. (2012). *Value Matters: Predictability of Stock Index Returns*. *SSRN Electronic Journal*, 1–27. https://ssrn.com/abstract=2031406

Arnott, R., Masturzo, J., & Kalesnik, V. (2018). *Cape fear: Why cape naysayers are wrong*. Retrieved May 14, 2023, from https://tinyurl.com/2p96kshb

Asness, C. S. (2002). *Fight the Fed Model: The Relationship between Stock Market Yields, Bond Market Yields, and Future Returns*. *SSRN Electronic Journal*, 1–33. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=381480

Baele, L., Bekaert, G., & Inghelbrecht, K. (2009). *The determinants of stock and bond return comovements* (Working Paper No. 15260). National Bureau of Economic Research. https://doi.org/10.3386/w15260

Begg, D., Vernasca, G., Fischer, S., & Dornbusch, R. (2014). *Economics* (11th ed.).

Berk, J., & DeMarzo, P. (2017). *Corporate finance: The core* (4th ed.). Pearson.

Brixton, A., Brooks, J., Hecht, P., Ilmanen, A., Maloney, T., & McQuinn, N. (2023). *A changing stock–bond correlation: Drivers and implications*. *The Journal of Portfolio Management Multi-Asset Special Issue*, *49*(4), 64–80. https://www.pm-research.com/content/iijpormgmt/49/4/64

Brooks, C. (2008). *Introductory econometrics for finance* (4th ed.).

Bunn, O., & Shiller, R. (2014). *Changing Times, Changing Values: A Historical Analysis of Sectors within the Us Stock Market 1872-2013* [https://ssrn.com/abstract=2478534]. *SSRN Electronic Journal*, 1–76.

Campbell, J., & Shiller, R. (2001). *Valuation ratios and the long-run stock market outlook: An update* (Working Paper No. 8221). National Bureau of Economic Research. https://doi.org/10.3386/w8221

Cochrane, J. (2005). *Asset pricing: Revised edition*. Princeton University Press.

Dickey, D., & Fuller, W. (1979). *Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association*, *74*(366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531

Enders, W. (2014). *Applied econometric time series* (4th ed.). Wiley.

Engle, R., & Granger, C. (1987). *Co-integration and error correction: Representation, estimation, and testing*. *Econometrica*, *55*(2), 251–276. Retrieved May 6, 2023, from http://www.jstor.org/stable/1913236

Estrada, J. (2006). *The Fed Model: The Bad, the Worse, and the Ugly*. *SSRN Electronic Journal*, 1–29. https://ssrn.com/abstract=877245

Grantham, J. (2017). *This time seems very, very different (part 2 of not with a bang but a whimper - a thought experiment)*. Retrieved May 25, 2023, from https://wealthtrack.com/wp-content/uploads/2018/01/This-Time-Seems-Very-Very-Different-by-Jeremy-Grantham.pdf

Hasseltoft, H. (2009). *The 'Fed Model' and the Changing Correlation of Stock and Bond Returns: An Equilibrium Approach*. *SSRN Electronic Journal*, 1–73. https://ssrn.com/abstract=1361489

Ilmanen, A. (2011). *Equity risk premium*. In *Expected returns* (pp. 117–151). John Wiley Sons, Ltd. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118467190.ch8

Jivraj, F., & Shiller, R. (2017). *The Many Colours of CAPE*. *SSRN Electronic Journal*, 1–25. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3258404

Johnson, N., Naik, V., Pedersen, N., & Sapra, S. (2013). *The stock-bond correlation*, 1–12. bit.ly/41wQpDw

Klement, J. (2012). *Does the Shiller-PE Work in Emerging Markets?* *SSRN Electronic Journal*, 1–41. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2088140

Koivu, M., Pennanen, T., & Ziemba, W. (2005). *Cointegration Analysis of the Fed Model. SSRN Electronic Journal*, 1–12. https://ssrn.com/abstract=685481

Li, L. (2002). *Macroeconomic Factors and the Correlation of Stock and Bond Returns. SSRN Electronic Journal*, 1–53. https://ssrn.com/abstract=363641

McDonald, R. (2013). *Derivatives markets*. Pearson Higher Education & Professional Group.

Micaletti, R. (2020). *Towards a Better Fed Model. SSRN Electronic Journal*, 1–33. https://ssrn.com/abstract=3704963

Modigliani, F., & Cohn, R. (1979). *Inflation, rational valuation and the market. Financial Analysts Journal*, *35*(2), 24–44. Retrieved May 6, 2023, from http://www.jstor.org/stable/4478223

Neville, H, Draaisma, T., Funnell, B., Campbell, R., & Hemert, O. V. (2021). *The Best Strategies for Inflationary Times. SSRN Electronic Journal*, 1–32. https://ssrn.com/abstract=3813202

Siegel, J. (2016). *The shiller cape ratio: A new look. Financial Analysts Journal*, *72*(3), 41–50. https://doi.org/10.2469/faj.v72.n3.1

Statsmodels. (n.d). Retrieved March 6, 2023, from https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html

Thau, A. (2010). *The bond book: Everything investors need to know about treasuries, municipals, gnmas, corporates, zeros, bond funds, money market funds, and more* (3rd ed.). McGraw Hill LLC.

Wooldridge, J. (2015). *Introductory econometrics: A modern approach* (6th ed.). South-Western Cengage Learning.

Zakamulin, V., & Hunnes, J. (2019). *Stock Earnings and Bond Yields in the US 1871-2016: The Story of a Changing Relationship. SSRN Electronic Journal*, 1–48. https://ssrn.com/abstract=3126321

NTNU

Norwegian University of
Science and Technology