Identifying tactical asset allocation signals for US and European equities

Valtteri Vaskikari

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.
Espoo June 14, 2020

Supervisor

Prof. Ahti Salo

Advisors

PhD Ruth Kaila

MSc Juhana Joensuu

Aalto University
School of Science

Aalto University
School of Science

Author Valtteri Vaskikari	
Title Identifying tactical asset allocation signals for US and European equities	
Degree programme Industrial engineering and management	
Major Strategy & Venturing	**Code of major** SCI3050
Supervisor Prof. Ahti Salo	
Advisors PhD Ruth Kaila, MSc Juhana Joensuu	
Date June 14, 2020 **Number of pages** 76	**Language** English

Abstract

Extensive research has been done to discover historically profitable tactical asset allocation strategies. Researchers have found a wide range of different market signals; valuation levels, company fundamentals, macroeconomic variables, sentiment indicators and seasonal patterns have been suggested to predict the market movements. However, it seems that most findings in published papers have been obtained with data-mining. Especially multiple tests have been performed without adjusting the significance level appropriately. In a typical journal article, predictors have also been tested only on in-sample periods, i.e., with the same data-set that was used for fitting the model.

The thesis re-evaluates the forecasting ability of the most potential stock market predictors found in the tactical asset allocation and equity market timing literature. The out-of-sample test results show that the equity premium has not been predictable in real-time after the turn of the millennium. The thesis thereby recommends the passive "buy and hold" strategy for both private and professional investors. However, if investors still want to try to "time the stock market" or actively adjust the strategic asset allocation mix, the inflation rate and ETF fund flows seem to be the most potential signals to follow.

Keywords Tactical asset allocation, Market timing, Stock market prediction, Type I error, Multiple testing, Bootstrapping, Data mining

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
Diplomityön tiivistelmä

Tekijä	Valtteri Vaskikari
Työn nimi	Taktisen allokaation signaalien tunnistaminen Yhdysvaltojen ja Euroopan osakemarkkinoilla
Koulutusohjelma	Tuotantotalous
Pääaine	Strategy and Venturing — **Pääaineen koodi** SCI3050
Työn valvoja	Prof. Ahti Salo
Työn ohjaaja	Ruth Kaila, Juhana Joensuu
Päivämäärä	June 14, 2020 — **Sivumäärä** 76 — **Kieli** Englanti

Tiivistelmä

Taktisen allokaation ja markkinoiden ajoittamisen tutkimukssa tutkijat ovat löytäneet laajalti erillaisia osakemarkkinoita ennustavia tekijöitä; arvostustasojen, yrityksen fundamenttien, makrotalouden muuttujien, sentimentti-indikaattoreiden ja kausivaihteluiden on havaittu ennustavan osakkeiden riskipreemiota. Vaikuttaa kuitenkin siltä, että suuri osa aikaisemmin julkaistuista löydöksistä on saatu tiedonlouhinnalla. Erityisesti useita tilastollisia testejä on tehty huomioimatta merkitsevyystason asianmukaisesta korjaamista. Useissa tutkimuksissa testit on lisäksi tehty samalla ajanjaksolla, jota käytettiin alkuperäisen mallin sovittamiseen.

Diplomityö arvioi uudelleen aikaisemmat taktisen allokaation ja markkinoiden ajoittamisen tutkimuksessa esitetyt löydökset osakemarkkinoiden ennustamisesta. Otoksen ulkopuolella tehdyt testit osoittavat, ettei osakeriskipreemioita ole ollut mahdollista ennustaa reaaliajassa vuosituhannen vaihteen jälkeen. Diplomityö siten suosittelee passiivista osta ja pidä –strategiaa sekä yksityisille että ammattimaisille sijoittajille. Mikäli sijoittajat kuitenkin haluavat yrittää taktista allokaatioita, potentiaalisimpia seurattavia markkinasignaaleita näyttävät olevan inflaatio ja ETF-pääomavirrat.

Avainsanat Taktinen allokaatio, Markkinoiden ajoittaminen, Osakemarkkinoiden ennustaminen, Tyypin I virhe, Moninkertaisen testauksen ongelma, Bootstrappaus, Tiedonlouhinta

Acknowledgement

Foremost, I am grateful for my thesis supervisor Ahti Salo. I have already greatly benefited from understanding the principles of good scientific writing. I also want to thank my advisor Ruth Kaila for the active and valuable guidance during the writing process. I would also like to express my gratitude to Juhana Joensuu and the rest of the quantitative analyst team of Elo for their active comments and feedback. I am also thankful for the finance professionals who gave the interviews and explained the complex causalities and phenomena related to the topic. Finally, I want to thank my friends and family for their support.

Otaniemi, 9.3.2020

Valtteri Elias Aleksander Vaskikari

Contents

Abstract 2

Abstract (in Finnish) 3

Acknowledgement 4

Contents 5

1 Introduction 10

1.1 Asset allocation . . . 10
1.2 Evaluating systematic TAA strategies . . . 10
1.3 Research setting . . . 11

2 Understanding the TAA signals presented in the literature 14

2.1 Market value . . . 14
2.1.1 Technical analysis . . . 14
2.1.2 Valuation ratios . . . 16
2.1.3 Fed Model . . . 18
2.2 Fundamental value . . . 20
2.2.1 Earnings announcement . . . 20
2.2.2 Payout ratio . . . 21
2.3 Real activity . . . 23
2.3.1 Productivity . . . 23
2.3.2 Inflation . . . 25
2.3.3 Interest rates . . . 26
2.3.4 Monetary base . . . 30
2.3.5 Credit . . . 31
2.4 World economy . . . 34
2.4.1 Exchange rates . . . 34
2.4.2 International trade . . . 35
2.4.3 Metal prices . . . 36
2.5 Market sentiment . . . 37
2.5.1 Sentiment surveys . . . 37
2.5.2 Implied volatility . . . 38
2.5.3 Put–call ratio . . . 39

2.5.4 Fund flow . . . 40
2.6 Other phenomena and anomalies . . . 40
2.6.1 Seasonal regularities . . . 40

3 Framework for empirical testing **42**
3.1 Statistical background . . . 42
3.1.1 Statistical hypothesis testing . . . 42
3.1.2 Multiple hypothesis tests . . . 43
3.1.3 Multiple testing correction . . . 44
3.2 Method for TAA model creation . . . 47
3.2.1 Initialization: panel regression and data transformations . . . 47
3.2.2 Stage 1: Orthogonalizing the X matrix . . . 49
3.2.3 Stage 2: Bootstrapping with replacement . . . 50
3.2.4 Stage 3: Signal selection . . . 51
3.3 Backtesting . . . 52
3.4 Test setting . . . 54

4 Results and discussion **55**

5 Conclusion **57**

References **59**

6 Appendix **64**

Symbols and abbreviations

A	Assets
a	Aggregate wealth
ACF	Autocorrelation function
AUM	Assets under management
B	Book value
BDI	Baltic Dry Index
BUS	Business
C	Private consumption, currency
c	Aggregate consumption
CAPE	Cyclically-adjusted price-to-earnings ratio,
CAY	Consumption to wealth ratio
CBO	Congressional Budget Office
CBOE	The Chicago Board Options Exchange
CPI	Consumer price index
Cr	Credit
CS	Credit spread
D	Dividend
DJIA	Dow Jones Industrial Average
DJTA	Dow Jones Transportation Average
D/P	Dividend yield
E	Earnings
EA	Euro area
ECB	European Central Bank
E/P	Earnings yield
EPS	Earnings per share
ERP	Equity risk premium
ETF	Exchange traded fund
FDR	False discovery rate
FED	Federal Reserve
FF	Fund flow
FOMO	Fear of missing out
FRED	Federal Reserve Economic Data
FWER	Family-wise error rate
FX	Forex, foreign exchange

G	Government spending and investment
g	Perpetual growth rate
GDP	Gross domestic product
GOV	Government
H	Hypothesis
HH	Household
I	Gross investment
i	Nominal interet rate
IMF	International monetary fund
LIBOR	London Interbank Offered Rate
M	Momentum, imports, money supply
m	Money multiplier, number of performed tests
MB	Monetary base
MSCI	Morgan Stanley Capital International
n	Sample size
NPO	Non-profit organization
NPV	Net present value
NXR	Nominal exchange rate
OLS	Ordinary leased squares
OMO	Open market operations
P	Price, probability
p	Probability value
PACF	Partial autocorrelation function
P/B	Price-to-book ratio
PCR	Put–call ratio
P/E	Price-to-earnings
PEAD	Post-earnings-announcement drift
QE	Quantitative easing
R	Return
R^2	Coefficient of determination
r	Real interest rate
RPR	Rolling percentile rank
SAA	Strategic asset allocation
SDW	ECB Statistical Data Warehouse
SMA	Simple moving average
S&P	Standard & Poor's

SR	Sharpe ratio
SRA	Sale and repurchase agreement
SUE	Standardized unexpected earnings
TAA	Tactical asset allocation
US	United States
VIX	Volatility Index
X	Exports
Y	Bond yield
y	Aggregate income
YC	Yield curve
YTM	Yield to maturity
α	Significance level, type I error rate
β	Regression coefficient, type II error rate
μ	Mean
π	Inflation
σ	Standard deviation
σ^2	Variance

1 Introduction

1.1 Asset allocation

Diversification is an essential principle of investment management. **The modern portfolio theory** suggests that investors obtain better risk-return trade-off by holding a broadly diversified portfolio rather than just a few investments (Markowitz, 1952). Diversification is based on the notion that adding more investments reduces the exposure to individual asset risk, and thereby decreases the overall risk in the portfolio. However, not all investments are equal in terms of diversification benefits. Investments within the same asset class, have, by definition, similar characteristics and are therefore subject to several same value drivers and risks. Thus, diversifying just within one asset class results in an inefficient portfolio, which does not offer the optimal return relative to the born risk. Therefore, investors need to diversify also across asset classes. (Sharpe, 1992)

Asset allocation determines the asset class weights in an investment portfolio. Research has shown that asset allocation is as important determinant of portfolio return as active management (Ibbotson, 2010). Asset allocation is divided into long-term **strategic asset allocation** (SAA) and short-term **tactical asset allocation** (TAA). SAA represents the mix of assets which is expected to meet investors financial objectives in the long-term, typically over a one-year horizon. Investors' overall risk-return profile, as well as diversification benefits, are the main factors affecting the strategic allocation decision. TAA, in turn, refers to the practice of actively adjusting the strategic asset allocation mix. TAA attempts to increase portfolio returns by over- and under-weighting asset classes according to temporary changes in their expected returns. **Market timing** is an extreme form of TAA that involves frequent shifts into and out of asset classes in an attempt to time the market peaks and troughs (Evensky et al., 2011). Over time, however, the strategic allocation is the most important determinant of the total return of a portfolio, while well-implemented TAA can contribute at the margin (Stockton & Shtekhman, 2010).

1.2 Evaluating systematic TAA strategies

TAA strategies can be either **discretionary** or **systematic**. In discretionary TAA strategies, allocation decisions are made based on the investment manager's judgement. In systematic TAA strategies, in turn, allocation decisions are based on a quantitative forecasting model. A **TAA model** uses financial and economic variables, or **signals**, to predict short-term asset class returns. Predictive signals are referred to as **bull** or **bear** (trade) signals according to their expected correlation with future returns.

Practitioners have found some best practices for evaluating systematic TAA strategies. Most importantly, strategies should be verified through robust empirical research. A strategy should generate significant excess returns whether it is tested on **in-sample** or **out-of-sample** periods. In-sample period refers to the data that is used for the initial parameter estimation and to the model selection. Parameters obtained from in-sample estimation would not have been known to investors in real-time, i.e., until the end of the sample. Therefore, findings should also be applied to an out-of-sample period, which refers to a data-set other than the one used for fitting the model.

However, "correlation does not imply causation" is a necessary disclaimer for all statistical tests. The phrase refers to a logical fallacy, in which two events occurring together are immediately considered to have an established cause-and-effect relationship (see "Super Bowl Indicator and Equity Markets" by (Schmidt & Clayton, 2017)). Thus, strategy evaluation should begin with disregarding all signals which conflict with economic intuition. All tested signals should have rational and logical explanations for their expected predictive power. Investment managers should also understand the underlying theories related to the signals and be aware of signals mutual connections. Moreover, investment managers need to be confirmed that empirical results are not obtained just by "mining the data", which refers, for example, to rerunning tests with modified signals until the statistically significant results are obtained.

Finally, systematic TAA strategies, like any other strategy, should be evaluated in terms of risk-adjusted returns. Strategies, which provide excess returns at the expense of a disproportionate increase in risk, are not actually value-adding because, through the use of leverage or leverage-like techniques, investors can always increase returns at the expense of increased risk. (Stockton & Shtekhman, 2010)

1.3 Research setting

The objective of the thesis is to identify tactical asset allocation signals for US and European equities. The objective is guided through three research questions:

1. Which signals are the most promising equity market predictors based on previous financial research and expert interviews?

2. Are the selected signals able to predict the time series of equity risk premium both in-sample and out-of-sample?

3. Are the results statistically significant when the effects of multiple testing and data mining are accounted for?

The first research question aims to review the vast equity market timing and TAA literature comprehensively. Experienced asset managers and investment bankers from Helsinki, London and Paris also contribute by sharing their knowledge and understanding of the topic. Companies, institutions and their representatives, however, prefer to remain anonymous. Nevertheless, in order to be selected for the empirical tests, logical reasoning is required to explain the signals' expected ability to predict market returns. Signals should also occur frequently because robust empirical testing requires multiple occurrences. Therefore, "fat tails" and "black swan" type of extremely rare events are excluded from the thesis.

The second research question makes an important distinction. The thesis concentrates exclusively on **time series** predictability in stead of **cross-sectional** predictability. Time series predictability refers to variation in the expected return for a given asset class over time. Cross-sectional predictability, in turn, refers to variation in the expected returns of various different assets at a given point in time. The second research question also guarantees that the signals are tested on out-of-sample periods as well.

Finally, the third research question addresses the severe problem of **data mining**. Data mining refers to all statistical processes which misuse sample data with the intention to enhance reported results. For example, performing multiple tests increases the likelihood that some variables exceed the predetermined significance level just by luck. In general, the more inferences are made, the more likely false inferences occur. Different disciplines use different terminologies for this problem. In physics, multiple testing is referred to as the "look-elsewhere effect". In medical science, and particularly in genome association studies, "multiple comparisons" is often used. In finance, "data mining", "data snooping", "overfitting", "p-hacking", "selection bias" and "**multiple testing**" are often used interchangeably.

The thesis is organized as follows. The second section presents the rationales and theories behind the most potential TAA and market timing signals. The third section introduces the empirical testing framework, which focuses on mitigating data-mining and especially the multiple testing problem. The fourth section shows the results, and also discusses the implications. Finally, the fifth section offers the concluding remarks and suggests further research areas.

2 Understanding the TAA signals presented in the literature

2.1 Market value

2.1.1 Technical analysis

Equity market prediction has long traditions in finance. Already at the end of 19$^{\text{th}}$ century, the founder of The Wall Street Journal and Dow Jones indexes, Charles Dow (1851 – 1902), developed a series of principles for interpreting and analyzing stock market behaviour. The principles became later known as the **Dow theory**, which is commonly regarded as the groundwork of modern **technical analysis**. As a part of his principles, Dow presented the oldest market timing strategy that is still used today. Dow suggested that the market would be bullish when both the Dow Jones Industrial Average (DJIA) and the Dow Jones Transportation Average (DJTA) reach historically high levels. Similarly, the market would be bearish when both indexes fall to historically low levels. If one would not follow the other, the movement was not likely the beginning of a new sustainable trend. (Schannep, 2008)

Dow assumed that in healthy business conditions, the growth of the industrial sector should also affect transportation services because airlines, railroads, shipping and other carriers would be required for the industrial activity. Thus, in order to detect long-term stock market trends, DJIA and DJTA should be analysed together (Schannep, 2008). Both indexes are still published today by The Wall Street Journal, and the Dow theory is typically mentioned in the financial news when the indexes reach new highs. Some investors still believe that the transportation stocks are a barometer of economic activity, and any market surge without their backing cannot be a long-lasting one.

Since the Dow Theory, a countless number of different technical indicators, methods and strategies have been developed to predict the stock market based on current and historical prices (P). Probably most popular are the **trend following** strategies, such as the **simple moving average** (SMA) which maintains a bullish signal as long as the current index price remains above the average of last n prices, i.e.,

$$SMA_n = \frac{P_t + P_{t-1} + ... + P_{t-n}}{n}. \tag{1}$$

The **breadth indicators** are another well-known group used in modern technical analysis. These indicators measure the number of advancing and declining stocks in the market to determine how many different shares are driving the index returns. Some investors use breadth indicators also as a measure of investor sentiment (see section 2.5). Nevertheless, most modern technical indicators are lacking both theoretical and empirical support. For example, Fang et al. (2014) re-examined the predictability of 93 previously presented technical indicators and found that most of them had none predictability, and even the best predictors could not beat the passive buy and hold strategy, i.e., the index return.

However, researchers still seem to agree that **momentum** (M_t) strategies may have some potential for tactical timing (Ilmanen, 2012). Momentum measures the rate of change in price over a predefined period of time, typically over past three to twelve months. The general form of momentum is usually presented as

$$M_t = \log\left(\frac{P_t}{P_{t-lag}}\right) * 100\%. \tag{2}$$

Researchers have found that most asset classes exhibit positive short-term momentum, and thereby investors can profit just by buying assets when the prices have risen and selling the assets when the prices have fallen. Researchers have found that momentum strategies can be profitably traded both across assets, by buying recent "winners" and selling recent "losers", but also on a single asset tactically over- and under-weighting the asset in the portfolio. (Ilmanen, 2012)

Researchers have explained momentum with theories from both rational and irrational finance. The latter is also known as **behavioral finance**. Some researcher believe that momentum is caused by institutional investors' slow portfolio rebalancing and decision making (Ilmanen, 2012). The market adapts new information slowly because significant shifts in institutions' portfolios require several meetings. Another theory is **the disposition effect**, which refers to investors tendency to sell too early the assets that have increased in value and hold too long the assets that have decreased in value (Shefrin & Statman, 1985). As a result, asset prices rise and decline more slowly than they should.

Yet another explanation is the **representativeness heuristic**, or "**apophenia**", which, in turn, refers to people's tendency to seek and see patterns in random information (Ilmanen, 2012; Kahneman & Tversky, 1979; Kahneman et al., 1982). Due to the representativeness heuristic, investors subconsciously want to extrapolate the observed trend even further, and buy (sell) securities that have recently increased (decreased) in value. Finally, momentum is also explained with investors **overconfidence bias**. Overconfident investors tend to undervalue public information and overvalue their own analysis, causing slow price reactions. Other explanations have included, stop-loss orders, margin calls and portfolio insurance strategies, i.e., hedging, among others (Antoniou & Koutmos, 2008; Ilmanen, 2012; Daniel et al., 1998).

2.1.2 Valuation ratios

Valuation ratios, such as the dividend yield (D/P), the earnings yield (E/P), the price-to-book ratio (P/B) and their different variants, are the most intuitive market timing indicators. It seems reasonable to assume that prices are not likely to drift too far away from their normal relationships with fundamental value, such as dividends or earnings. Therefore, it seems intuitive that when valuation ratios are very high by historical standards, prices would fall in the future and restore the normal levels. Thus, it seems reasonable that there would be a relationship between current valuations ratios and subsequent market returns. Figure 1 seems to confirm the intuition by implying that lowest valuation ratios would produce the highest returns.

However, Asness and Ilmanen (2017) show that outperforming a passive buy and hold approach with valuation signals is harder than Figure 1 would suggest. They show that valuation ratios can drift higher or lower for years or even decades, making it surprisingly difficult to categorize the current market confidently as "cheap" or "expensive" in real-time. Asness and Ilmanen found that valuation ratios generally drifted lower during the early 1900s than during the last 60 years. This upward drift means that this timing strategy would have got excessively many "sell" signals in recent decades. For example, their test strategy started to obtain a "sell" signal already in 1991 rather than 1999 or 2000 when the dot-com bubble finally burst. This strategy thereby lost almost a decade of profits from equity risk premium (ERP), demonstrating very well the fundamental difficulty with all TAA strategies: "too early equals wrong".

Asness and Ilmanen also tested the hypothesis that valuation signals would work better when applied only at extreme valuations. They show that even this enhancement does not significantly improve their results. Finally, they try to exploit

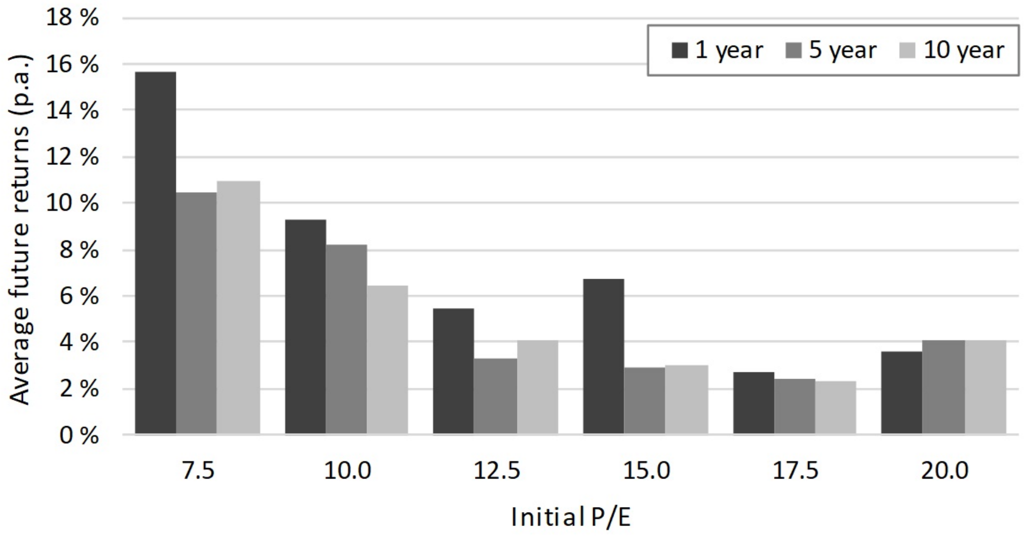

Figure 1: Future returns are higher when market valuations are cheap, S&P500 1871-2019 (Source: Robert Shiller's website). Figure is constructed by sorting each month of S&P 500 data into nearest P/E group (x-axis), and then computing the average returns for one, five and ten years ahead (y-axis). Only groups of relevant size are presented, and thereby groups after P/E 20 and below 7.5 are omitted.

the drifting valuation phenomenon by combining the valuation signal with momentum. The strategy that used an average of value and momentum signals gave better results than both signals separately. Asness and Ilmanen argue that combining valuation ratios and momentum mitigates the risks of **value traps**, i.e., further drops, and too early "sell" signals.

Another explanation for the poor predictability of valuation ratios was already given by Graham et al. (1934). Graham et al. had argued that most financial ratios were too sensitive to medium-term business cycles and other events, and therefore were not able to depict firms' sustainable value. Graham et al. argued that the current prices should rather be compared to long-term averages of fundamental value. Professor Robert Shiller popularized this idea by using cyclically-adjusted price-to-earnings (CAPE) ratio

$$CAPE = \frac{P}{E_{\text{real 10y avg.}}}, \tag{3}$$

to smooth cyclical variations in earnings, and successfully predicted the dot-com bubble in 2000 and the US housing crash seven years later.

Other researchers have found that the **dividend yield** would have the best timing ability among different valuation ratios. Though up to the beginning of the 1980s, buybacks were minimal. As a result, data for this signal is only available 1980s, whereas other valuation ratios can be tested from the 19th century. Some practitioners, in turn, prefer value timing strategies that combine the main valuation ratios into one combined measure (Ilmanen, 2011).

2.1.3 Fed Model

The Federal Reserve (FED) model is an asset allocation analysis comparing earnings yields and bond yields to determine their relative attractiveness. The Fed model states that the fixed-income market and the stock market are in equilibrium when the stock market's one-year forward-looking earnings yield equals the 10-year Treasury note yield, i.e.,

$$\frac{E_{1y,\ forward}}{P} = Y_{10y,\ GOV}. \tag{4}$$

The Fed model equilibrium can be derived from the Gordon growth model formula

$$P = \frac{D(1+g)}{R_f + ERP - g}, \tag{5}$$

assuming that 100% of the earnings are paid as a dividend (D), the perpetual growth rate (g) equals zero, the 10-year Treasury note yield equals the risk-free rate (R_f) and there is no equity risk premium (ERP). The Fed model suggests that if markets deviate from the equilibrium, the higher-yielding asset should be over-weighted in the portfolio.

While the Fed model has been acknowledged to have several drawbacks, it still has its supporters (Asness, 2002). Most practitioners regard that the basic comparison between bonds and stocks is valid. Practitioners argue that in practice, stocks and bonds are competing assets in investors' portfolios. If one yields significantly better, the funds start to flow, and the demand increases for the other. On a large scale, this should eventually return the model's equilibrium. Another argument supporting the model concerns the relationship with stock price valuation and the risk-free rate. Asset pricing theory suggests that the value of stocks should be equal to the sum of its discounted future cash flows. So, if the risk-free rate drops, the discounted cash flows will be greater, leading to higher stock price, and thus to lower earnings yield. As the government bond rate is commonly regarded as a proxy for the risk-free rate, this implies that both sides of the equation would then drop by the same amount, sustaining the equilibrium in the model. (Campbell & Vuolteenaho, 2004)

Despite questionable model assumptions, researchers have found that the Fed model been able to time the markets. Asness (2002) explained these findings with a cognitive bias known as **money illusion**. This bias describes the confusion with nominal and real, i.e., inflation-adjusted, values. This confusion occurs easily while using the Fed model because the left-hand side (E_{1y}/P) is presented in real terms, whereas the right-hand side ($Y_{10y,\ GOV}$) is a nominal value. This is caused by the different approaches in stock and bond pricing. Inflation should not move stock prices because future cash flows rise by an amount that offsets the increase in the discount rate (see section 2.3.2). The increase in one-year forward-looking earnings is negligible, leaving the left-hand side of the Fed model nonreactive to changes in inflation. Bond prices, in turn, move in the opposite direction with inflation because coupons to bondholders are fixed, so the change in the discount rate is not offset. As a consequence, increasing inflation increases the bond yield (as bond prices drop) whereas earnings yield remains still. So, Asness proposes that when sufficiently many investors exhibit money illusion and compare real earnings yield to nominal bond yields, rising inflation tends to make stocks undervalued (high E_{1y}/P) while falling inflation tends to make stocks overvalued (low E_{1y}/P).

There are many reasons why even finance professionals are prone to confuse nominal and real values. A common explanation is that people regard that nominal presentation is just simpler, and usually sufficient first approximation of real value. It is considerably easier and more natural to think values in nominal terms rather than in real terms. People also perceive that this mistake is small, and the difference between the two can be neglected at least when inflation is low and stable. Researchers have also found evidence that the money illusion diminishes in hyperinflation because people start to draw attention to inflation more seriously. (Ilmanen, 2011)

Even though empirical findings strongly support the existence of money illusion, practitioners have still found difficulties to exploit this anomaly profitably. Money illusion tends to be slow-moving, affecting years or even decades at a time. It is also difficult to diversify as it tends to affect across asset classes. Its appearance is also inconsistent. Researchers have found that money illusion can explain the low valuations during the 1970s and high valuations in the 1990s. However, it still fails to explain the cheap valuations in the deflationary 1930s and the boom of the housing market amidst high inflation in 1970s. (Ilmanen, 2011)

Other practitioners have explained the Fed model's timing ability with **reflexivity**

theory. Already in 1987, George Soros suggested that market prices do not just reflect the expected future outcomes but also can change them (Soros, 2015). He claims that financial markets operate in "forward-looking bias" that can validate itself by influencing the fundamentals that market prices should reflect. This theory originates from sociology, where a concept **self-fulfilling prophecy** refers to the phenomenon where "prediction" or expectation comes true simply because people believe it will. So, some practitioners believe that the Fed model has influenced markets because of its popularity. In favour of this theory, it is known that influential market participants, such as the former chairman of FED, Alan Greenspan, has used it. (Greenspan, 2008)

2.2 Fundamental value

2.2.1 Earnings announcement

Post-earnings-announcement drift (PEAD) refers to the phenomenon where stock prices continue moving gradually after the earnings announcement. PEAD was first discovered by Ball and Brown (1968) who found that after a sharp move on the earnings announcement day, stock prices tended to drift upwards for several months if the reported results had exceeded investors' expectations. In accounting and financial research, this phenomenon is usually measured with the **standardized unexpected earnings** (SUE). SUE is the difference between the stock market's aggregated earnings per share (EPS) and the last forecast. Usually SUE standardized by dividing the measure by its historical standard deviation, i.e.,

$$SUE = \frac{EPS - EPS_{forecast}}{\sigma(EPS - EPS_{forecast})}. \tag{6}$$

Several theories have been proposed to explain the PEAD. Most widely accepted view is that the market under-reacts to earnings announcement because of **anchoring**. In cognitive psychology, anchoring is a cognitive bias where people depend too heavily on the initial information, "the anchor", when making subsequent judgments. Investors exhibit anchoring bias when they try to adjust their initial analysis with the new additional market information. The subsequent adjustments are typically insufficient, leaving the new assessments biased towards the initial analysis. **Confirmation bias** is another cognitive bias usually related to anchoring. It refers to people's tendency to search for, interpret, favour, and recall new information in a way that it affirms the prior view. These theories then suggest that the PEAD occur because investors are too attached to their original analyses and are slightly stubborn to change their view.

2.2.2 Payout ratio

In a series of papers published since 1979, Robert Shiller proposed that investing is essentially a social activity. Shiller et al. (1984) discovered that investors spend a significant share of their leisure time with conversations about investing, reading about investments or gossiping about others success and failures in investing. They proposed that, like conversations about any other popular topic, such as food, clothing and politics, change people's attitudes, discussions about investing shape also investors' attitudes about their investments. They found that the attitudes, preferences and fashions are usually shared widely among the population, and often appear without any rational reason.

Shiller et al. challenged the efficient market hypothesis by suggesting that asset price movements are mainly driven by psychology, and particularly, by fashions and fads. Shiller et al. argued that fashions and fads cause excess variability in the markets amplifying market reactions. They claimed that some of the people's psychological reactions are predictable, and thus stock markets could be predicted as well, at least to some extent. Especially, they found that the stock market seemed to overreact to dividends. According to their findings, historically high dividends seemed to cause over-optimism in the market and increase prices in the short-term. Afterwards, prices tended to return towards their long-term trends.

Increased dividend is typically considered as an indication of positive future prospects because managers and the board of directors should have the best information about the firm's future. Investors assume that excess cash is distributed to owners if the management anticipates the following years to be profitable so that that future cash flows can cover the upcoming capital expenditure. This theory is also known as **dividend signalling**. While Shiller et al. did not challenge the ideas of dividend signalling, they still found the market reactions were unjustified; the market should not have reacted to dividends as much as it did. They explained this overreaction both with social elements, such as fads and fashions, and behavioural biases. Shiller et al. cited particularly Tversky and Kahneman (1974) who had found that people tend to overreact to small probability events, and under-react to large probabilities (currently an element of the **prospect theory**). Shiller et al. then suggested that the prospect theory explains the dividend overreaction because surveys of corporations' dividend policy had found that managers try to keep dividends fairly constant through time. Thereby, the deviations in dividend policy is, in this sense, a low probability event (Lintner, 1956).

In the late 1990s, Lamont (1998) continued the earlier research of Shiller et al. and found that aggregated dividend payout ratio (D/E) was able to predict short-term stock market returns. Lamont explained the results with **Lintner's dividend policy model**, which assumes that dividends are paid according to managements forward-looking payout ratio target. Lamont referred to earlier work of Lintner and explained that firms tend to set forward-looking targets for payout ratios according to the amount of NPV positive projects they currently have available. Managers acknowledge that increased earnings are not always sustainable, and thus dividend policy is not changed until new earnings levels can be maintained consistently. Thereby, like dividends alone, increased payout ratio signals that management foresees more positive future prospects.

2.3 Real activity

2.3.1 Productivity

Real activity refers to the economy of a **single market**, or "internal market". A single market is an economic area in which most trade barriers have been removed and where people, goods, services and capital can move freely. Researchers have found that real activity and stock market returns are linked in complicated ways. To begin with, researchers have been interested in the role of economic productivity, which is most commonly measured with the **gross domestic product** (GDP). GDP represents the final value of the goods and services produced within an economy during a specified period of time, usually a year. GDP equals the total sum of private consumption (C), gross investment (I), government spending and investment (G), and the balance of trade (X-M), i.e.,

$$GDP = C + I + G + (X - M). \tag{7}$$

Due to the definition of GDP, it seems natural to expect that changes in GDP would affect companies' sales and other financials, such as earnings, and henceforth to dividends and market returns.

However, empirical research has shown that GDP and stock market's aggregated EPS are surprisingly weakly related. Ilmanen explains that new start-ups, ventures and other unlisted companies often captures the most GDP growth in the economy. Arnott and Bernstein (2002) have also suggested that EPS is not affected by GDP because of equity dilution. Existing firms have tended to issue new shares effectively along with the economic growth. Due to these weak relations, researchers have found that GDP expectations and GDP growth do not have significant market return predictability. (Ilmanen, 2011)

Nonetheless, researchers have found that other productivity measures may add value for TAA. Based on previous research, one potential measure is the **GDP gap**, or "output gap". GDP gap measures the difference between actual and potential GDP, and it is usually presented in percentages from the potential GDP, i.e.,

$$gap_t = \frac{GDP_{actual} - GDP_{potential}}{GDP_{potential}} * 100\%. \tag{8}$$

Potential GDP, or economy's "potential", is the estimated level of production that an economy should be able to sustain over long-term without negative consequences. This level depends on supply-side factors of the economy, such as the supply of workers and their productivity. Potential GDP corresponds to **full employment**, which means that anyone willing and able to work at the prevailing rate of wages, is employed. Full employment thereby means that there is no involuntary unemployment, but some frictional, structural and voluntary unemployment may still exist.

So, if the GDP gap is negative, an economy is producing at a level that has "slack". In this situation, the economy is below the full-employment, and thus some labour, capital, or other resources are underutilized. An economy may also have positive GDP gap, which means that an economy operates above its full employment. Positive GDP gap may occur when the economy's aggregate demand exceeds the level of demand required to establish full employment. In this case, the shortage in labour increases the wages, which, in turn, start to increase prices of products and services, causing an overheating economy and inflation. Thus, both negative and positive GDP gap is regarded as harmful to the economy. Several researchers have suggested that the negative effects on real activity explain why both positive and negative GDP gap has predicted a bearish stock market. (Cooper & Priestley, 2008; Vivian & Wohar, 2013; Ahmad & Sharma, 2018)

Few productivity-related indicators have also been suggested for TAA. For example, Lettau and Ludvigson (2001) reported predictive relations from **consumption to wealth ratio**, or "CAY", which dives the consumers' aggregate consumption (c_t) by aggregate wealth (a_t) and aggregate income (y_t), i.e.,

$$cay_t = \frac{c_t}{\beta_a a_t + \beta_y y_t}. \tag{9}$$

Lettau and Ludvigson explained the findings with a theory where consumers want to maintain stable consumption over time with respect to their asset wealth and labour income. Lettau and Ludvigson suggested that, when forward-looking consumers expect higher future total earnings from wages, investments, and other sources, they react by raising their consumption in advance. Consumers thereby allow their consumption to rise above its common relationship to the wealth. Similarly, when consumers expect lower personal income, they decrease their spending in advance. Thus, consumption to wealth ratio signals consumers' future income and spending, which then affects to real activity, firms' sales and stock prices. However, the results of Lettau and Ludvigson were later questioned because the regression coefficients β_a and β_y were based on in-sample fit, i.e., using the whole time period data, implying that the values of cay_t would not have been available to investors in real-time.

2.3.2 Inflation

Inflation is defined as the increase of general price level, reducing the purchasing power each unit of currency (C) can buy. Inflation is typically caused by overheating economy (see section 2.3.1) but several other causes also exist. The basic economic theory between inflation and stock returns was established by Fisher (1930). The fundamental **Fisher equation** is

$$i \approx r + \pi, \tag{10}$$

and it estimates the relationship between nominal (i) and real (r) interest rates under inflation (π). The equation states that the nominal interest rate is approximately equal to the sum of the real interest rate and inflation. **Fisher hypothesis**, or the "Fisher effect", then follows by stating that real assets should be hedged against inflation because nominal asset returns move one to one with the discount rate (see section 2.1.3). Thereby, real stock returns should be independent of inflation because normal (common) stock represents ownership of the income generated by real assets. Asset pricing theory also suggests that stock prices should not be affected because higher inflation and discount rate will be fully compensated by higher nominal earnings growth. However, empirical research has shown that the Fisher hypothesis does not hold in practice.

In 1980s, Fama (1981) found a negative relationship between inflation and stock returns. Fama proposed that this was induced by negative relations between high inflation and real activity. Fama's ideas became later known as the **proxy hypothesis**. Fama described several reasons why high inflation is typically regarded as harmful to the overall economy, and why these harm firms as well. For example, **menu cost** are the costs to firms resulting from constantly reconsidering and changing their prices. High inflation also increases economic uncertainty which makes firms reluctant to make investments because managers are uncertain about future operational costs, personnel expenses and consumer demand. Postponing capital expenditures then has several negative consequences, such as falling behind international competitors.

According to Bekaert and Wang (2010) the negative relation between stock returns and inflation can be explained with **stagflation**. Stagflation is an economic situation combining stagnant or low economic growth, high unemployment, and high inflation. Stagflation is a difficult condition because the main actions of monetary and fiscal policies can not fight both unemployment and inflation at the same time. Policies aimed at lowering inflation in the long-term will typically increase unemployment in the short-term. In turn, policies aimed at lowering unemployment in the short-term translates to even higher inflation in the long-term. Stagflation can occur when both **expansionary**, i.e., stimulating, and **contractionary**, i.e., hindering, policies co-occur. For example, if both the central bank expands the money supply, and the government increases taxes simultaneously, inflation increases even though economic growth starts to slow. Stagflation is detrimental for firms and stocks because companies face at the same time rising costs but also lower demand.

Other researchers have found that in certain circumstances, a positive relationship between inflation and stock returns may occur. Main explanations have included deflationary environment, the money illusion (see section 2.1.3), and distortions in reported earnings due to accounting and taxation rules. There are a few different ways how the accounting and taxation rules may seemingly boost the reported earnings. For example, **inventory profit** may occur if the nominal value of the inventory happens to increase before the inventory is sold. Earnings can also increase if depreciation becomes insufficient in real terms due to unexpected inflation.

To summarize, all previously presented explanations may contribute to the complicated relationship between inflation and the stock market. Despite mixed results and somewhat contradicting theories, most researchers seem still to agree that steady, low, but slightly positive inflation, such as 2%, would be the optimal environment for the real activity and the stock returns. Policymakers and central bankers also seem to support this view because 2% is the targeted inflation rate in most developed economies. (Ilmanen, 2011)

2.3.3 Interest rates

Interest rate is the amount of money a lender charges for a loan. Interest rates are usually expressed as a percentage of the original borrowed sum, which is also known as the **principal** or the **face value**. Interest rates can also refer to the **yield to maturity** (YTM), which is the annual return earned on a bond if the bond is bought at the current market price and held until it matures. An essential graphical presentation of interest rates is the **yield curve** (YC), which shows the YTMs of bonds that share the same credit quality but have different maturities. Yield curve steepness is usually measured as the interest rate difference between a long-term, such as 10-year, and a short-term, such as 2-year, government bonds, and thereby the general form can be written as

$$YC = Y_{\text{long-term}} - Y_{\text{short-term}}. \tag{11}$$

Most of the time, two main factors are affecting the steepness of the yield curve. The monetary policy of the central bank determines the short-term rates, while the market's inflation expectations determine the long-term rates. (Ilmanen, 2012)

Manipulating the front end of the yield curve is the central bank's primary mean to control inflation. Most frequently, the central bank manipulates the short-term interest rates by raising and lowering the target for the **overnight rate**. The overnight rate is the interest rate at which commercial banks and the other financial institutions can lend to each other overnight on an uncollateralized basis. Thus, the overnight rate is also known as the **interbank rate**. Manipulating this rate is essential for the central bank because it influences indirectly to other interest rates as well, such as mortgage loans, and thus in some extent steers the economy. Typically, a monetary policy-making body of the central bank decides the target rate a couple of times in a quarter. Inflation, unemployment, economic growth and the overall state of the economy are the main drivers affecting the decision. (Dalio, 2012)

However, the central bank cannot force commercial banks to charge that exact target rate, because the effective overnight rate is determined through negotiations between the commercial banks. The effective federal funds rate is then calculated as the weighted average of interest rates across all actual overnight lending transactions. Therefore, the central bank typically attempts to affect the overnight rate indirectly. As the overnight rate is usually the lowest available interest rate and only available to the most creditworthy institutions, it is highly influenced by the yield of the government bonds. So, the central bank attempts to affect the overnight rate by manipulating the short-term government bond yields. By buying and selling government bonds in the open market, the prices and thereby the yields can be affected. As in general, buying bonds in the open market increases the price and lowers the yield, and correspondingly, selling bonds decreases the price and increases the yield. Central bank can also manipulate the overnight rate by entering into a **sale and repurchase agreement** (SRA). In SRA, the central bank lends money to commercial banks for overnight basis and requires assets, such as government bonds, as collateral. Both mechanisms are jointly referred to as the central bank's **open market operations** (OMO). (Dalio, 2012)

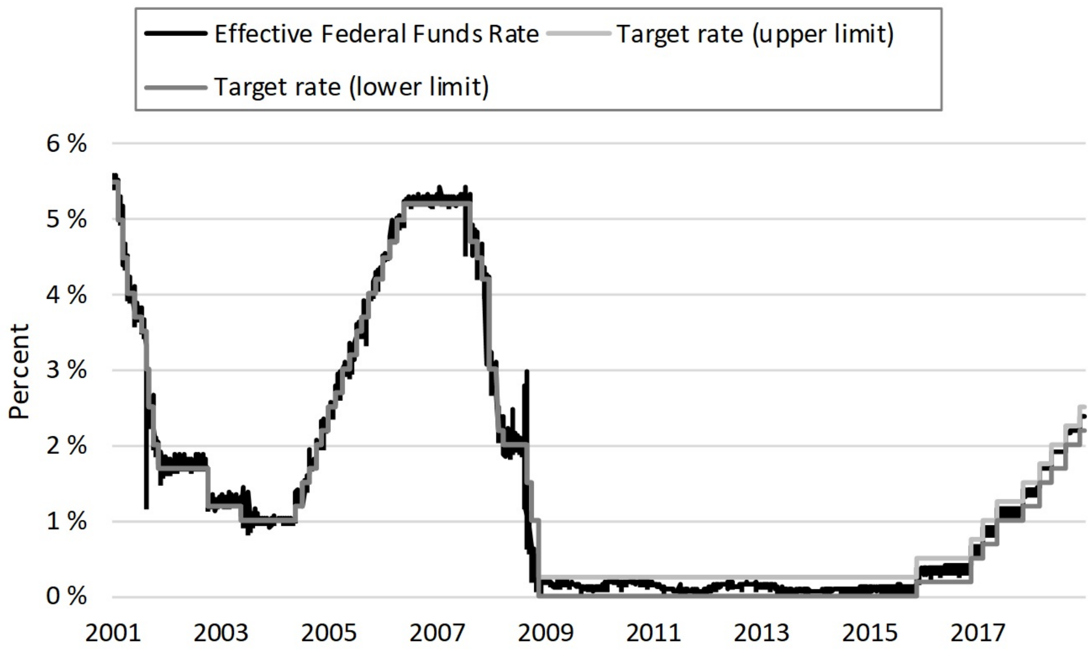

Figure 2: Open market operations have historically been rather successful to obtain the desired overnight rate in the US. (Source: Board of Governors of the Federal Reserve System, US. Upper and lower limits reported separately since 16.12.2008)

The central bank also has other tools to influence to the front-end of the curve. **The discount rate** is the interest rate charged to commercial banks and other financial institutions on loans they receive from the "**the discount window**", which is the central bank's lending facility. The discount rate is usually higher than the overnight target rate, that encourages banks to borrow from each other and only turn to the central bank if necessary. **Reserve requirements** are, in turn, the portions of deposits that commercial banks must hold either in their own vaults or as deposits at the central bank. The historical purpose of reserve requirements is to prevent "**bank runs**", which refer to a situation where bank's clients start to withdraw their money in fear that the bank ceases to function in the near future. In general, the central bank executes expansionary monetary policy by purchasing bonds through OMO, lowering the discount rate and decreasing the reserve requirement. In contrast, the central bank executes contractionary monetary policy by selling bonds through OMO, increasing the discount rate and increasing the reserve requirements. OMO, the discount rate and the reserve requirements are collectively referred to as the three **instruments of monetary policy**, and the three mechanisms are referred to as the **conventional policy measures**. (Dalio, 2018)

Central bank's conventional policy measures do not directly affect the back-end levels of the yield curve. Instead, the long-term rates are mainly driven by the bond investor's inflation expectations. The yield curve is upward sloping when the central bank executes expansionary monetary policy, and the bond market expects high inflation in the future. Long-term debt has a higher risk, and thus the bond market rationally compensates the risk with a higher yield. However, the yield curve flattens or even inverts when the central bank starts to execute contractionary policy measures, and the bond market starts to anticipate low inflation. The inversion is rational because the inflation-adjusted yield of the long-term debt still remains higher than the yield of the short-term debt as the inflation is expected to be lower in the future. Though, inflation expectations are not solely the only factor affecting the inversion. Some investors also believe that the long-term government bonds, such as U.S. 10-year Treasury bond, serve as **safe haven**. Investors' safe haven refers to assets that are expected to retain, or even gain value during periods of economic downturn. (Ilmanen, 2011)

An inverted yield curve has been believed to predict a recession for a few reasons. Firstly, contractionary policies aim to hinder economic growth. Secondly, low inflation is most often associated with recessions. And thirdly, increased demand for safe haven assets signs economic uncertainty. Researchers have also found that the stock market crashes before the beginning of a recession because the prices start to reflect the forthcoming difficulties. Researchers have also suggested that the stock market crash follows the yield curve inversion, and thereby the stock market crash could be foreseen. Researchers have explained that yield curve inverts before the stock market crashes because the bond market is "smarter" than the stock market. The **smart money** theory suggests that most bond market participants are professional investors rather than amateurs, and therefore the bond market should reflect future expectations more accurately. (Harvey, 1989)

Researchers have also found other relationships between the interest rates and stock prices. The most direct relationship is the **present value effect**; lowering interest rates increase the stock prices because of a milder discount factor. Researchers have also proposed that an inverted yield curve can also be a self-fulfilling prophesy (see section 2.1.3). This theory suggests that firms frighten the yield curve inversion and its threatening signal of recession. As a consequence, companies start to postpone their new investments and recruitments, which then starts to slow the economic growth and eventually causes the turn of the business cycle by itself. (Harvey, 1989)

2.3.4 Monetary base

Most central banks have a dual mandate to maintain both stable price level and sustainable economic growth. Most of the time, the conventional monetary policies (see section 2.3.3) are sufficient to steer the economy through the business cycle in a way that the dual mandate is fulfilled. Occasionally, however, the economic conditions become so extreme that the conventional measures become ineffective. Such situations include **inflationary depression**, i.e. stagflation (see section 2.3.2) and **deflationary depression**. (Dalio, 2012)

Deflationary depression is more common between the two. It is defined as an economic situation combining low economic growth and low inflation. In such a situation, despite conventional measures lowers the short-term rates to zero, or the "**zero bound**", the economy still requires further stimulation to recover. Deflationary depression is difficult because the central bank can not further lower short-term interest rates due to **liquidity trap**. Liquidity trap refers to a situation where investors start to prefer cash over debt securities because there is no opportunity cost of holding cash if the nominal interest rate is zero. Therefore, the central bank will have to resort to **unconventional policy measures**, which include, for example, **quantitative easing** (QE) and **helicopter money**. Helicopter money, or "helicopter drop", refers to rather rare monetary policy where the central bank distributes money directly to citizens in a certain form of dividend. (Dalio, 2012)

QE, in turn, has occurred frequently. In QE, the central bank creates electronic money, and buys government bonds and other riskier financial assets from private financial institutions, or "**depository institutions**". The large-scale asset purchases inject money into the banking system more than OMO (see section 2.3.3), and thus, QE is assumed to be more effective than conventional policy measures. In QE, the total central bank assets (A_{total}) increase. Particularly, a balance sheet construct known as the **monetary base** (MB), or the "central bank money", grows. Depository institutions' cash funds over reserve requirements also rise. These funds are known as **excess liquidity** or **excess reserves** (ER). So, the central bank desires that QE would ensure low short- and long-term interest rates, i.e., flatten the yield curve. A low interest rate environment should then encourage businesses and people to borrow and consume more. Another aim is to help the government to manage through the recession. In recessions, the government' income tax usually drops due to higher unemployment, while the spending on social benefits increases. QE eases the government to finance the budget deficits by increasing national debt rather than just increasing taxes. (Dalio, 2012)

QE affects asset prices through several channels. In **credit channel**, lower interest rates decrease firms and households' cost of borrowing, raise corporations capital expenditure, and thereby raise asset prices through recovering economy. However, typically QE programs carry no restrictions for depositary institutions on how the new central bank money should be used. Commercial banks are not obliged to lend the money to the households, and thereby the additional liquidity is not necessarily passed on to the non-financial sector. Instead, typically the institutions that have sold assets in QE prefer to rebalance the portfolios with listed financial assets rather than consumer loans. So, another known channel is the **portfolio rebalancing**. Rebalancing process inflates across all asset classes because the central bank has withdrawn a substantial amount of safe haven assets from the market, and thereby investors need to turn to other riskier assets as well. (Gambetti & Musso, 2017) Researchers have found that asset classes which are considered riskier than bonds, such as equities, appreciate gradually over a multiple week window following the QE announcement days. Mamaysky (2014) explained that in most funds, allocating investments to riskier assets requires more time because consent from senior managers is usually required (see momentum theories in section 2.1.1). Finally, QE announcements have also a **signalling effect** which refers to central bank's view on the forthcoming economic conditions. (Dalio, 2018)

2.3.5 Credit

In **fractional reserve banking** only a proportion of commercial bank deposits are backed by actual cash and available for withdrawal. Rest of the deposited capital is freed for lending. Commercial banks can then create additional "commercial bank money", or "**credit**" (Cr), by loaning a proportional amount of each deposit in the bank. Each granted loan becomes a new deposit in the same or another bank, and a proportional amount of this new deposit can be then lent further. The established loop is known as the **multiplier effect**. Due to the multiplier effect, the initial central bank money will be lent and deposited multiple times resulting that the actual money supply, i.e., the money in circulation, will be a multiple of the original monetary base. The theoretical maximum for the money supply at any given point of time is expressed with the **money multiplier** (m):

$$M = m \times MB. \tag{12}$$

The money multiplier is defined as the inverse of the capital reserve ratio, which is the minimum fraction of the total deposits that commercial banks have to maintain. (Dalio, 2018)

The money supply is measured with **monetary aggregates**. Monetary aggregates categorize the different forms of money, depending on the ease to convert an asset or holding into actual cash. **M1** aggregate (M_1), for example, includes the most liquid portions of the money supply, such as physical currency and deposits held in the commercial banks. The total amount of credit in the economy, or the "debt burden", is, in turn, usually presented as a sum of the total debt taken by the government (GOV), businesses (BUS), households (HH) and non-profit organizations (NPO). Debt burden is usually expressed as a percentage of total economic output, i.e.,

$$\text{Debt burden} = \frac{D_{\text{GOV}} + D_{\text{BUS}} + D_{\text{HH}} + D_{\text{NPO}}}{GDP} * 100\%. \tag{13}$$

Theories of debt, credit and leverage cycles explain the relationship between the economy's total debt and asset prices. The debt cycle theories suggest that the total debt is the most significant driver of the business cycle. However, contrary to the common belief, debt is not fundamentally bad for the economy. Most economists actually believe that debt plays an essential role in allocating resources effectively. Debt also establishes projects, investments and other opportunities that increase economic productivity in the long-term. Economists have even argued that too low debt burden can create as bad or even worse economic problems as having too much debt. (Dalio, 2012)

However, researchers suggest that market bubbles and debt crises emerge because the total debt starts to increase faster than the economic productivity. Bubbles begin to grow when loans are granted too easily, and taken projects do not increase economic productivity by sufficient amount, i.e., have a negative NPV. Typically, reckless lending begins during the periods of how high economic growth. Reckless lending has been explained with **herd-behaviour** and **availability bias**. The latter refers to people's tendency to weigh recent experience more heavily than would be appropriate. Asset prices soar because the created credit starts to increase economic activity, but also because new investors enter the financial markets. New investors start to enter the markets due to social anxiety known as the **fear of missing out** (FOMO). FOMO is defined as a compulsive concern, fear or regret of missing a potentially rewarding opportunity by not participating in a certain activity. Debt cycle theories then suggest that increased asset prices then enable further lending because the worth of collaterals, such as houses, have risen along with the asset prices. As the investors' confidence increases, lending starts to increase also in less regulated "**shadow banking**" sector. Shadow banking system refers to credit intermediation involving entities and **financial engineering** activities outside of the regulated banking system. (Geanakoplos, 2010)

In debt cycle theories, the bust of the debt bubble is then triggered by the central bank's contractionary monetary policy and especially, the rise of short-term interest rates. Due to inflation pressures, the central bank starts to increase the short-term interest rates, causing the yield curve to invert (see section 2.3.3). As credit becomes more expensive, consumption starts to decrease, which, in turn, decreases people's and businesses' income. Also, debt repayments become more challenging because borrowing more money to repay the existing loans has become more expensive. As a result, investors rush to sell their assets, causing asset prices to fall. Due to the falling asset prices, the worth of collaterals drop as well, and thereby debt defaults start to cause severe losses to lending institutions. The vicious cycle is then established, eventually leading to a debt crisis, stock market crash, and to a **deleveraging** process where the debt levels within the economy are drastically reduced.

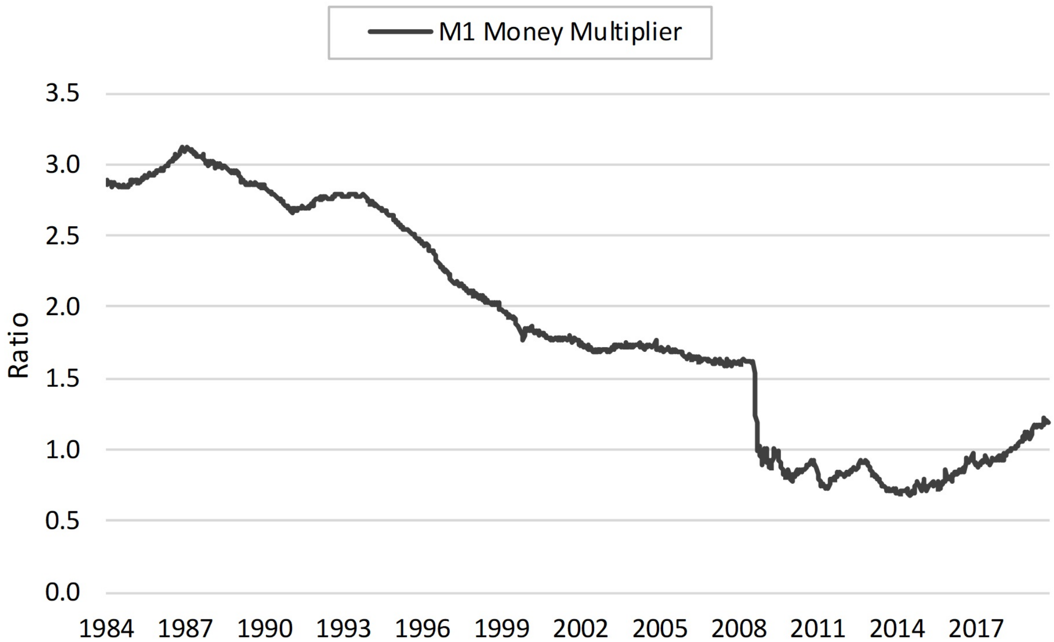

Figure 3: M1 money multiplier is the ratio between M1 monetary aggregate and the monetary base. The multiplier dropped when deleveraging started in the financial crisis of 2008 (Source: Board of Governors of the Federal Reserve System, US)

Researchers and practitioners have tried to tactically time the burst of the debt bubble by monitoring the actions of the biggest lenders and credit institutions. Chava et al. (2010) has examined the lending standards of Federal Reserve, and Adrian et al., in turn, has examined the balance sheets of **financial intermediaries**, such as shadow banking institutions. Some researchers have also followed the **credit impulse**

$$\frac{\mathrm{d}^2(Cr/GDP)}{\mathrm{d}t^2} = \frac{Cr_t - Cr_{t-1}}{GDP_t} - \frac{Cr_{t-1} - Cr_{t-2}}{GDP_{t-1}}, \tag{14}$$

which measures the change in newly issued credit (Cr). Practitioners have also monitored the **credit spread**, which is the difference between low- and high-grade corporate bonds, i.e.,

$$CS = Y_{\text{Low-grade}} - Y_{\text{High-grade}}. \tag{15}$$

2.4 World economy

2.4.1 Exchange rates

Exchange rate is defined as a price at which one currency will be exchanged for another. An exchange rate can be presented in relation to one another currency or in relative to a "**basket**" of multiple foreign currencies. Most exchange rates are free-floating, and thus are determined by supply and demand in the **foreign exchange** (FX) market. The prices in the FX market are highly influenced by real activity. For example, if all else equal, inflation depreciates the currency because it decreases the purchasing power. Thereby, changes in the money supply affect as well because money supply affect inflation (see section 2.3.4). Increasing interest rates, in turn, appreciate the domestic currency because higher interest rates provide higher rates to lenders, thereby attracting more foreign capital to the economy. However, the debt burden and negative balance of trade, i.e., the trade deficit, typically depreciate the currency due to increased defaulting risk and economic uncertainty. Researchers have also found a variety of other factors driving the FX market, including geopolitical events, terms of international trade, among others. (Sekmen, 2011)

However, because the FX market is a "zero-sum game" with relative prices, all currencies do not react to changes in the global macroeconomic envrionment in the same way. For example, while most currencies depreciate, some currencies are regarded as safe haven assets and tend to appreciate systematically during negative events. A currency can become a safe haven asset in a few different ways. For example, if the central bank maintains continuously low interest rates, that currency may become attractive for **currency carry** trading. In a currency carry trade, FX traders attempt to profit from the difference in the interest rates between two currencies. FX traders borrow the currency with the low interest rate, exchange that to the other currency, and then invest with the higher interest rate. Currency carry trading is profitable if the exchange rate does not change and offset the interest rate difference. The safe haven property of a currency can then be explained with FX traders rush to close the currency carry positions when global uncertainty increases, leading to a rapid peak in demand and price during negative macroeconomic events. (Yau & Nieh, 2006)

Yau and Nieh (2006) have found that safe haven currencies show early warning signs of bearish stock markets. The main explanation is that, like the bond market, the FX market is considered "smarter" than the stock market (see section 2.3.3). Even though retail investors are a growing segment in the FX market, their trading volume is nonexistent compared to professional investors. The most common strategies, such as carry trades, also require high leverage, which is usually unreachable for most amateur investors. In addition, professional investors in the FX markets have significantly better access to the data, and they have excessive resources to research the market continuously. Thus, exchange rates are mainly moved by professional investors, and the FX market is thereby considered rather efficient.

However, whether or not varying exchange rate favours the domestic stock market, depends on the industrial structure within the economy. Stocks benefit from the currency depreciation if the economy is export-driven. As a result of deprecation, exporters improve their competitiveness in the international markets, resulting in higher sales and profits. (Yau & Nieh, 2006). However, excessively volatile currency is harmful for stocks because hedging the foreign exchange risk becomes more expensive, and typically hedging instruments do not completely lessen the lost trade volume. (Sekmen, 2011)

2.4.2 International trade

The Baltic Dry Index (BDI) is the most followed indicator of international trade activity. The BDI is a weighted average of international shipbrokers' assessments of shipping costs with the **Capesize**, **Panamax** and **Supramax** dry bulk transport vessels. The three dry bulk vessel types differ on their weight carrying capacity and size. Vessel size determines the vessel's ability to travel through different maritime trade canals, such as the Panama Canal. Therefore, different vessels have significantly different shipping costs and voyages, i.e., routes. The assessments used in the BDI, are given daily for 20 internationally essential shipping routes: 5 Capesize, 5 Panamax and 10 Supramax voyages. The selected voyages are meant to have large enough a trade volume to represent the global maritime freight transport. In the BDI, the shipping costs are measured as **time chartering** (TC) averages, which include fuel costs, port charges, commissions, and a daily crew hire. The BDI formula is then written as

$$BDI = \frac{0.4 * TC_{\text{avg. 5 Capesize}} + 0.3 * TC_{\text{avg. 5 Panamax}} + 0.3 * TC_{\text{avg. 10 Supramax}}}{10}. \tag{16}$$

The supply of dry bulk cargo ships is rather tight and inelastic because it takes several years to build a new vessel. Therefore, the BDI is mainly driven by the demand for dry bulk, which includes mainly industrial raw materials, such as coal, iron ore, steel, cement, and grain. The BDI is thereby considered as a forward-looking indicator of global industrial activity. For this reason, movements in the BDI have a significant influence on political and economic decision-making. Industrial companies also use the BDI for both operative and strategic decisions. One of the main reasons for the popularity of the BDI is that governments, associations or investors cannot influence or manipulate it easily. However, financial researchers have found that the growth rate of BDI is able to predict stock market trends in the short-term. Researchers have explained that in most economies, the BDI predicts the development of manufacturing industry that will then drive the other sectors and henceforth asset prices as well. (Bakshi et al., 2010; Apergis & Payne, 2013)

2.4.3 Metal prices

The interpretation of industrial metal prices is analogous to the Baltic Dry Index signal. Industrial metals are, by definition, raw materials for the industrial sector, and thereby the prices reflect the demand for global industrial production and construction. However, unlike the BDI, metal prices are also influenced by supply-side factors, such as the global metal production capacity, which weakens the relationship with the stock market (Jacobsen et al., 2018). In contrast, precious metals, and especially gold, have similarities with the safe haven currencies (see section 2.4.1). Gold is commonly seen as a reliable safe haven asset because it has a historical legacy as a store of value, currency and wealth. Researchers have also suggested that gold's bright, shiny and positive image possibly contributes to investors preference for gold during economic downturns (Huang & Kilic, 2019; Baek, 2019). However, researchers have found only mild relationships between the gold price and stock market returns. (Erb & Harvey, 2013)

2.5 Market sentiment

2.5.1 Sentiment surveys

Market sentiment, or "investor sentiment", is defined as investors overall feeling and attitude towards the market. Various measures and indicators have been developed in order to approximate the market sentiment. Probably the simplest indicators of investor sentiment are the sentiment surveys where investors answer to questionnaires in regular time intervals. Typically, sentiment surveys are conducted weekly, and investors choose among three alternatives: "bullish", "bearish" or "neutral", depending on their view on the stock market over the following months. Sentiment surveys have been arranged at least since 1989 when Rober Shiller started to conduct them at Yale University as a part of his research (Brown & Cliff, 2005). However, the results of the surveys are typically published in the form of **bull–bear ratio**

$$\text{Bull–bear ratio} = \frac{\text{Bullish}}{\text{Bullish+Bearish}} * 100\%, \tag{17}$$

which accounts only the bullish and bearish investors, and ignores the neutral respondents.

Researchers have found a negative relationship between "bullish" survey results and stock market returns. Thus, the bull–bear ratio is a **contrarian indicator** and thereby encourages to trade in contrast to the prevailing sentiment. According to Stambaugh et al. (2012), investor sentiment may increase prices excessively due to **overoptimism** and herd-behaviour. Stambaugh et al. have explained that such overvaluation may then prevail even for a long period of time, because most investors are either restricted or otherwise unwilling to short-sale. Stambaugh et al. found that some professional investors are prohibited from selling short, and most private investors, in turn, find short selling riskier or otherwise uncomfortable. Thus, most investors tend to take no position at all rather than sell short if they believe that the market is overvalued. Stambaugh et al. therefore suggest that during periods of high market-wide sentiment, the market tends to be slightly overpriced and future returns will thereby be low. Researchers have found that such mispricing does not occur when the survey results are natural or bearish level.

2.5.2 Implied volatility

Since introduced in 1993, The Chicago Board Options Exchange (CBOE) Volatility Index (VIX) has been considered as a premier gauge of investor sentiment. VIX measures the implied volatility of 30-day S&P 500 index options derived from the Black-Scholes formula. According to option pricing theory, higher volatility of the underlying asset makes both put and call options more valuable, because there will be a higher probability that the option will expire in the money. Thus, high measures of VIX imply that investors expect a sharp upward or downward movement in the market. However, as the market crashes only downwards, the implied volatility accounts more the downside risk. For this reason, the VIX is also known as the "fear index", and associated with bearish market expectations. The volatility index can be calculated with the formula

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\triangle K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2, \tag{18}$$

where
VIX $= \sigma * 100$
T = Time to expiration
F = Forward index level derived from index option prices
K_0 = First strike below the forward index level, F
K_i = Strike price of i^{th} out-of-the-money option
$\triangle K_i$ = Interval between strike prices
R = Risk-free interest rate to expiration
$Q(K_i)$ = The midpoint of the bid-ask spread for each option with strike K_i

Several theories have been proposed to explain the negative relationship between VIX and the stock market. For example, Brunnermeier and Pedersen (2008) suggest that higher implied volatility makes **market-makers**, or "liquidity providers", more conservative and less active in facilitating the transactions in the market. Declined market-making decreases the market liquidity and thereby also prices. Adrian and Shin (2010), in turn, suggest that liquidity and prices decline because financial intermediaries, such as brokers and commercial banks, have certain preventive risk management mechanisms which start to constrain lending in times of high VIX. Other researchers have also proposed that hedge funds tend to lose some of their assets under management (AUM) in times of high implied volatility. In such times, hedge funds typically reduce the leverage as well. Hedge funds' decreasing AUM and leverage thereby cause more selling in the market and put downward pressure on the prices. (Nagel, 2012)

2.5.3 Put–call ratio

The put–call ratio (PCR) is another popular indicator of investor sentiment. The ratio represents a proportion between all purchased put and call index options on a given day. The PCR is presented either in volume or dollar-weighted basis, i.e., for example

$$PCR = \frac{\text{Put volume}}{\text{Call volume}} * 100\%. \tag{19}$$

As put options are used to hedge against falling prices, a large proportion of puts to calls indicates bearish investor sentiment. However, most investors use the PCR as a contrarian indicator (see section 2.5.1). Contrarian investors believe that the volume of sold put and call options should be roughly equal over time. Thus, deviations from the historical average imply that the market may be turning excessively bullish or bearish. Thus, for a contrarian investor, a large proportion of puts or calls indicates excessive sentiment and potential market overreaction. The contrarian investors have also argued that PCR may reveal the positions of the institutional investors because most institutional investors need to hedge their positions. For example, a large proportion of puts may then imply that institutional investors have long positions which they just have hedged, causing the high PCR. Thus, contrarian investors also follow the "smart money" (see sections 2.3.3 and 2.4.1). (Martikainen & Puttonen, 1996)

2.5.4 Fund flow

Fund flow (FF) represent the net amount of new money invested in a fund in a given period of time. The general form can be written as

$$FF = \frac{\text{Inflow} - \text{Outflow}}{\text{Net Assets}} * 100\%. \tag{20}$$

The fund flow of exchange-traded funds (ETF) is typically interpreted as a measure of investor sentiment: positive ETF flow indicates that investors expect a bullish market and negative fund flow correspondingly implies that investor are feeling bearish. However, empirical research has found that ETF flow predicts just the opposite market reaction: future returns are high when ETF flow is low, and future returns are low when ETF flow is high. Thus, like the bull–bear and put–call ratios (see section 2.5.1 and 2.5.3), ETF flow is regarded as a contrarian indicator. (Frazzini & Lamont, 2008)

The negative relationship between ETF flow and the stock market has been explained with the **dumb money effect**. The dumb money effect suggests that amateur investors, rather than professionals, drive ETF flows. Researchers have explained that ETF funds predict low returns because most behavioural biases, such as excessive optimism or pessimism, affect especially amateur investors. Amateur investors are also known to be the least informed segment of investors, and therefore they are most likely to buy and sell at the wrong time. Another proposed explanation is that certain hedge funds tend to switch their holdings to ETFs when they anticipate uncertain times in the market. Hedge funds find some ETFs more liquid than their underlying assets, and thereby ETFs are less risky alternative during a bear market. Thus, increased ETF flow may reflect both amateur investors' excessive sentiment, but also professional investors' preventive actions. (Tabs, 2010)

2.6 Other phenomena and anomalies

2.6.1 Seasonal regularities

Higher average returns in January is probably the best-known seasonal anomaly. The January effect is usually explained with a rebound from an earlier drop in December. Researchers have suggested that market declines in December because investors try to reduce their capital gain taxes by selling poorly performed investments at a loss before the year's end. The January effect has also been explained with corporations annual bonus payments which occur mostly in the end of December or in the beginning of January. Some of the bonus money is then used for stock purchases, causing increasing stock prices.

Another well-known anomaly is the Halloween effect which, in turn, refers to lower returns in May–October. This regularity is usually explained with summer holidays. Research suggests that investors are generally more risk-averse during a summer vacation. Most financial professionals, such as equity research analyst and investment managers, are also less active or entirely out of offices during the summer months.

In recent years, end of month effect has attracted attention due to the publication of Etula et al. (2015). They found that payments of pensions funds, dividends of corporate treasuries, and distributions of mutual funds, all occur at the turn of a month. Etula et al. show that in every month, financial assets worth billions of US dollars are liquidated and distributed to pensioners and investors before the month-end. The money is then partially re-invested at the beginning of the next month, causing a regular return-pattern around the turn of every month. In addition to the January, Halloween and end of month effects, researchers have also found higher average returns around major holidays and macroeconomic news announcements.

3 Framework for empirical testing

3.1 Statistical background

3.1.1 Statistical hypothesis testing

Statistical hypothesis testing is a method of statistical inference that is used to interpret and draw conclusions about a population using sample data. A hypothesis test evaluates two mutually exclusive statements: the **null hypothesis** (H_0) and the **alternative hypothesis** (H_1). The null hypothesis generally states that there is no significant association between variables, or that there is an insignificant relationship between measured phenomena. The null hypothesis often has a form "there is no difference". Alternative hypothesis is contrary to the null hypothesis, suggesting that a statistically significant relationship exists. The alternative hypothesis, or the "**research hypothesis**", is often the logical opposite of the null hypothesis, and thus usually has a form "there is a difference". The null hypothesis is assumed to be true until statistical evidence indicates otherwise, in which case it is rejected. If the null hypothesis is rejected, the alternative hypothesis is accepted.

Sample's deviation from the null hypothesis is measured with a **test statistic**. The probabilities for different test statistic values are obtained from the **null distribution**, which is the probability distribution given that the null hypothesis is correct. The null distribution value associated with the observed test-statistic value is denoted as the **probability value** (p) By the definition of the null distribution, p-value measures the probability of obtaining similar or more extreme test statistic value just by chance, given that the null hypothesis is true. Therefore, the smaller the p-value, the stronger the statistical evidence that the null hypothesis should be rejected. The null hypothesis is rejected, and results are considered **statistically significant** if the p-value is less than or equal to the pre-defined **significance level** (α), i.e.,

$$p \leq \alpha \rightarrow \text{ reject } H_0 \tag{21}$$

$$p \geq \alpha \rightarrow \text{ accept } H_0. \tag{22}$$

Two conceptual errors have an essential part in statistical hypothesis testing. A **type I error**, or "false positive" finding, is the rejection of a true null hypothesis, while a **type II error**, or "false negative" finding, is the non-rejection of a false null hypothesis. In simpler terms, type I error incorrectly finds a relationship between variables although there is none, while type II error fails to detect a relationship that really exists. The type I error rate (α) is the probability of making a type I error in a test. The type I error rate is at most as high as the predefined significance level.

Similarly, the type II error rate (β) is the probability of making a type II error in a test. It is, in turn, related to the **power** of a test, which is the probability that the test correctly rejects the null hypothesis when the alternative hypothesis is true (see table 2). In most disciplines type I errors are considered more severe than type II errors. This originates from the traditions in legal trials, where type II error occurs when a guilty person is set free and remain unpunished. This is considered far less serious error than type I error where an innocent person is convicted and suffers the sentence.

	True H_0	False H_0
Accepted H_0	Correct inference (true negative) Confidence level $1-\alpha$	Type II error (false negative) β
Rejected H_0	Type I error (false positive) Significance level α	Correct inference (true positive) Power $1-\beta$

Table 2: Type I and type II errors

3.1.2 Multiple hypothesis tests

The multiple testing problem arises when a statistical analysis involves multiple hypothesis tests, each having a potential to produce statistically significant results, i.e., a "discovery". Multiple testing is problematic because significance levels that apply to each test individually are rarely sufficient for the whole set, or "**family**", of simultaneous tests. Unless the performed tests are perfectly positively dependent, i.e., identical, the probability to make at least one false discovery increases rapidly along with the number of performed tests. Therefore, the significance level requires an adjustment to compensate for the multiple comparisons. In general, the more simultaneous tests are performed, the more stringent significance level is required.

However, the number of performed test is not the only determinant for the appropriate significance level correction. The similarity, or the "cross-sectional dependency" between the tests, affects as well. For example, if the performed tests are perfectly independent, then the probability of making one or more false discoveries is

$$
\begin{aligned}
P(\textit{Type I error}) &= \alpha \\
P(\textit{Not making a type I error}) &= 1-\alpha \\
P(\textit{Not making a type I error in m tests}) &= (1-\alpha)^m \\
P(\textit{Making one or more type I errors in m tests}) &= 1-(1-\alpha)^m.
\end{aligned}
$$

In this case, suppose we have, for example, 100 independent tests, and each test has a significance level of 5%. The chance to make at least one false discovery is $1-(1-0.05)^{100} \approx 99.4\%$

Another important special case of multiple hypothesis testing is a set of perfectly dependent tests. In this case, the same test is just performed multiple times, and as a result, there is effectively only one hypothesis tested. So, in this case, the probability of type I errors does not change, and thereby there is no need to adjust the significance level. For example, suppose we have 100 identical tests, and each test has a significance level of 5%. The probability of making at least one type I error equals $1-(1-0.05)^1 = 5\%$

3.1.3 Multiple testing correction

A variety of different approaches has been developed to adjust the level of significance for multiple comparisons. Two commonly used approaches for controlling type I errors are the **family-wise error rate** (FWER)

$$FWER = P(V \leq 1), \tag{23}$$

and the **false discovery rate** (FDR)

$$FDR = E\left[\frac{V}{R}\right]. \tag{24}$$

FWER is the probability of making at least one type I error when multiple hypotheses are tests. FDR, in turn, is the expected proportion of type I errors among the set of rejected null hypotheses. If R is the total amount of rejected null hypotheses, and m_0 is the number of true null hypotheses, FWER and FDR can be defined using the following table

	True H_0	False H_0	Total
Accepted H_0	U	T	$m - R$
Rejected H_0	V	S	R
Total	m_0	$m - m_0$	m

Table 3: Possible outcomes of multiple tests

Classical FWER control procedures include the Bonferroni correction (Bonferroni et al., 1936), the Šidák correction (Šidák, 1967), the Tukey's range test (Tukey et al., 1949), the Holm–Bonferroni method (Holm, 1979), the Hochberg's step-up procedure (Hochberg, 1988), the Dunnett's correction (Dunnett, 1955), the Scheffé's method (Scheffé, 1953) and the Harmonic mean p-value (Good, 1958). In turn, the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) and the Benjamini–Yekutieli procedure (Benjamini et al., 2001) are classical solutions for controlling the FDR.

Classical multiple testing controls have certain advantages and disadvantages. In general, FWER controls are well protected against type I errors, but the statistical power of FWER controls is usually weak, especially when the number of tests increases. Due to low power, there is a low probability that FWER control detects the real discoveries, i.e., correctly rejects the false null hypothesis. FDR controls, in turn, have high power, but it comes with the expense of an increased number of type I errors. For these reasons, FWER controls are preferred in applications where the number of tested hypotheses remains small, and when the consequences of a single false discovery are severe. FDR controls, in turn, are preferred in applications where the amount of hypotheses is excessively large, and when a small proportion of false discoveries is tolerable. Thus, economic and financial research usually favours FWER controls. In contrast, FDR controls are often appropriate, for example, in genetic epidemiology where relationships with genes and thousands of different diseases are under investigation. In genetic epidemiology, a small proportion of false discoveries is tolerable because these can be identified in subsequent studies.

That being said, financial researchers have found that most classical FWER controls are still too strict even with a low number of performed tests. Researchers have argued that the classical FWER controls too often detect nothing, i.e., nothing can be declared statistically significant in the presence of multiple testing. In most classical FWER controls, this "over-conservatism" is caused by the assumption of complete independence between the tests. Financial researchers have argued that almost all economic and financial variables are interlinked in complicated ways, and thereby the tests should not be considered independent. So, motivated by the shortcomings of the classical procedures, more sophisticated FWER controls have been developed in recent years. The goal of these methods has been to guard the type I errors without excessively sacrificing the statistical power. The most promising approaches have turned out to be the **resampling** methods, such as **permutation tests** and **bootstrapping**. (Westfall et al., 1993; Romano & Wolf, 2005)

However, financial researchers have regarded that the permutation tests are somewhat challenging due to their excessive computational requirements. Bootstrapping, in turn, has received support. In many recent financial publications, bootstrapping has been preferred because bootstrapping methods do not have to appeal to any asymptotic theories in order to provide inference. For example, an assumption about the shape of the null distribution, such as normality, is not required. Distribution assumptions are not required because bootstrapping preserves the dependency information in the data. Particularly, the statistical **moments** are maintained. Moments are the statistics obtained from a random sample, and they characterise the shape of the **sampling distribution**, which is the theoretical probability distribution of the population. The **mean** (μ) and **variance** (σ^2) are the first and the second moments of a sampling distribution, and **skewness** and **kurtosis** are the third and the fourth moments, respectively. Moments beyond 4$^{\text{th}}$ order are referred to as the **higher-order statistics** and they describe the non-linear dependencies in the data. As bootstrapping preserves both lower- and higher-order moments, the joint null distribution of the multiple tests can be approximated directly from the sample with the use of sampling distribution.

3.2 Method for TAA model creation

3.2.1 Initialization: panel regression and data transformations

The goal of the method is to create a TAA model that consists only statistically significant TAA signals. The method includes an algorithm that iterates three stages and adds a new signal to the model after each iteration. The three stages are repeated as long as new statistically significant signal can be found to improve the model incrementally. During each iteration, the confidence level is adjusted considering the total amount of candidate signals remaining and their mutual dependence (see section 3.1.2). The applied technique is a bootstrapping based FWER control procedure (see section 3.1.3).

So, suppose there is a $T \times 1$ vector Y of equity market returns that $T \times M$ candidate signal matrix X attempts to predict. Each column of this matrix X presents a time-series of a TAA signal's test variable, and vector Y presents the market return of the following period. The whole multi-dimensional data, or "**panel data**", can be presented in a matrix form where rows denote time periods and columns denote test variables

$$\begin{bmatrix} Y & X_1 & X_2 & \cdots & X_m \end{bmatrix} = \begin{bmatrix} y_2 & x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ y_3 & x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ y_4 & x_{3,1} & x_{3,2} & \cdots & x_{3,m} \\ y_5 & x_{4,1} & x_{4,2} & \cdots & x_{4,m} \\ y_6 & x_{5,1} & x_{5,2} & \cdots & x_{5,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_t & x_{t-1,1} & x_{t-1,2} & \cdots & x_{t-1,m} \end{bmatrix} \tag{25}$$

The method includes modelling relationships between X_i and Y. As typical for financial research, predictive relationships are modelled with linear regressions. The parameters of the regressions are estimated with the **the ordinary leased squares** (OLS) method, which minimizes the sum of squared residuals, i.e., the distances between the predictive model and data points. The goodness-of-fit is measured with the **coefficient of determination** (R^2), which describes the proportion of the variance that each X_i can explain. While linear regressions and the OLS method are widely used in a variety of different applications, there are still certain requirements which need to be fulfilled in ordered to justify their usage. The main general assumptions include:

- The modelled relationship is linear.

- The mean of the error terms is zero.
- The error term variance is constant over time (also known as the "**homoscedasticity** of residuals").
- The error term values are not **autocorrelated**.
- Independent variables X_i are uncorrelated with the error term.
- In multivariate regressions, no variable can be linearly approximated from the other variables, i.e., there is no **multi-collinearity**.

Certain **transformations**, are typically required for the data so that the assumptions of the linear regression and OLS method are met. Data transformations are mathematical functions which are applied to every data point in the time series. Through the use of data transformations, linear regressions can also model relatively complex relationships, such as nonlinear dependency, assuming that the appropriate transformations to linearity can be found. With linear regression models, it is usually sufficient to ensure that all used time series are, at least approximately, **stationary**, which means that the statistical properties, such as mean, variance, auto-correlation and the error term, are all constant over time. Visualizations, such as plots of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), are typically used to determine whether transformations have been able to obtain sufficient stationarity. Data transformations may also be applied in order to change the interpretability of the variables. Main transformations include:

- **Difference**: transforms each data point in the time series into difference between the current and the previous value

  $$z_t = x_t - x_{t-1}. \tag{26}$$

  Difference is typically applied for variables which are expected to affect through the changes in the absolute values.

- **Log difference**: converts each data point into the relative difference between the current and the previous value

  $$z_t = \log x_t - \log x_{t-1}. \tag{27}$$

Log difference is considered as an appropriate transformation for variables that are expected to function through the changes but do not have any natural reference point, such as zero. Log difference, thereby, reveals more information about the magnitude of the difference.

- **Z-score**: standardizes a time series to obtain a mean of zero and a standard deviation of one

$$z_t = \frac{x_t - \mu}{\sigma}. \tag{28}$$

Z-score is an in-sample measure because μ and σ are calculated using the whole sample. As a result, normal z-score values are not stable over time, and thereby would not have been available to investors in real-time. Another variant, the **rolling** z-score, uses only that data which would have been available at that given point of time, i.e., μ and σ are calculated for each z_t using only values $x_1, x_2...x_t$. However, as the data increases gradually, the very first rolling z-score values usually fluctuate aggressively due to very few used data points. Therefore, a certain **warmup perdiod** is usually required to remove the first rolling z-score values from the inference. Rolling z-score standardization is typically applied to variables that are assumed to affect through a deviation from the historical mean. Multiple variables may also be combined into one **composite** z-score by taking a mean of the individual z-score values. Composite z-score is used to gauge different variables combined effect.

3.2.2 Stage 1: Orthogonalizing the X matrix

The algorithm begins after the required data transformations have been applied. So, suppose that k $(0 \leq k \leq M)$ TAA signals have already been selected to the TAA model and there is $M - k$ candidate signals remaining. As $M - k$ is also the number of performed tests during each iteration. So, the first step is to **orthogonalize** the X matrix. In this context, orthogonalization refers to dividing the time series of each candidate signal into components and then removing that component which seems to have incremental predictability. Incremental predictability of the X matrix can be separated and removed with **vector projections**. Vector projections model a linear relationship between each candidate signal X_{k+j} and residual vector $Y^{e,k}$

$$X_{k+j} = c + dY^{e,k} + X^e_{k+j}, \quad j = 1, ..., M - k. \tag{29}$$

Here c an is the intercept, d is the slope and X^e_{k+j} denotes the error term of the projection. During the first iteration the $Y^{e,k}$ equals the Y vector because no signals have yet been selected to the model, and thereby model residuals are considered to be the original Y vector values. As the X^e_{k+j} is the error term of the projection, it should be independent and uncorrelated with the residuals of the current model. The matrix X can then be orthogonalized by replacing each candidate signal column with their corresponding X^e_{k+j} vectors. As a result, the new X matrix should not provide any incremental predictability, and thereby all individual regressions between candidate signals and model residuals should have zero R^2 values.

3.2.3 Stage 2: Bootstrapping with replacement

New orthogonalized data matrix (30) is shown below. Here $Y^{e,k}$ is the residual vector of the current model, X^s is the submatrix of preselected variables and X^e is an orthogonalized submatrix of candidate variables:

$$\begin{bmatrix} Y^{e,k} & X^s & X^e \end{bmatrix} = \begin{bmatrix} y_2^{e,k} & x_{1,1}^s & \cdots & x_{1,k}^s & x_{1,k+1}^e & \cdots & x_{1,m}^e \\ y_3^{e,k} & x_{2,1}^s & \cdots & x_{2,k}^s & x_{2,k+1}^e & \cdots & x_{2,m}^e \\ y_4^{e,k} & x_{3,1}^s & \cdots & x_{3,k}^s & x_{3,k+1}^e & \cdots & x_{3,m}^e \\ y_5^{e,k} & x_{4,1}^s & \cdots & x_{4,k}^s & x_{4,k+1}^e & \cdots & x_{4,m}^e \\ y_6^{e,k} & x_{5,1}^s & \cdots & x_{5,k}^s & x_{5,k+1}^e & \cdots & x_{5,m}^e \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ y_t^{e,k} & x_{t-1,1}^s & \cdots & x_{t-1,k}^s & x_{t-1,k+1}^e & \cdots & x_{t-1,m}^e \end{bmatrix} \tag{30}$$

Even though the orthogonalized X matrix should not be able to explain the model residuals, some statistical predictability may start to emerge if the matrix rows are **bootstrapped with replacement**. This technique constructs a new matrix of same size by mixing rows of the orthogonalized matrix (30) randomly, and in a way that an individual row can emerge multiple times. One bootstrap could look like the following:

$$\begin{bmatrix} Y^{e,k} & X^s & X^e \end{bmatrix} \xRightarrow{\text{Bootstrap}}$$

$$\begin{bmatrix} y_8^{e,k} & x_{7,1}^s & \cdots & x_{7,k}^s & x_{7,k+1}^e & \cdots & x_{7,m}^e \\ y_{13}^{e,k} & x_{12,1}^s & \cdots & x_{12,k}^s & x_{12,k+1}^e & \cdots & x_{12,m}^e \\ y_{41}^{e,k} & x_{40,1}^s & \cdots & x_{40,k}^s & x_{40,k+1}^e & \cdots & x_{40,m}^e \\ y_8^{e,k} & x_{7,1}^s & \cdots & x_{7,k}^s & x_{7,k+1}^e & \cdots & x_{7,m}^e \\ y_{116}^{e,k} & x_{115,1}^s & \cdots & x_{115,k}^s & x_{115,k+1}^e & \cdots & x_{115,m}^e \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{13}^{e,k} & x_{12,1}^s & \cdots & x_{12,k}^s & x_{12,k+1}^e & \cdots & x_{12,m}^e \end{bmatrix} = \begin{bmatrix} Y_{boot}^{e,k} & X_{boot}^s & X_{boot}^e \end{bmatrix} \tag{31}$$

After bootstrapping, $M-k$ regressions are made between X_{boot}^e and $Y_{boot}^{e,k}$. As rows were bootstrapped with replacement, the R^2 values can deviate even significantly from zero. So, the highest R^2 value is saved among the $M-k$ regressions, and then this step is repeated 10 000 times. After a large number of repeats, the saved R^2 values can be ordered and presented as an empirical distribution. This distribution approximates the joint null hypothesis that none of the $M-k$ candidate signals can incrementally improve the model. Thereby, the empirical distribution represents the null distribution of the multiple tests.

3.2.4 Stage 3: Signal selection

Once the null distribution has been obtained, the signal selection process follows the steps of standard statistical hypothesis testing (see section 3.1.1). Test statistics, i.e., the "real" R^2 values, for each TAA signal can be obtained by performing individual regressions between $Y^{e,k}$ and each X_i from the original panel data (25). The p-values for these test statistic values can then be obtained from the approximated null distribution.

During each iteration, the null hypothesis is that none of the remaining candidate signals have additional explanatory power to predict the future equity market returns. The null hypothesis is rejected, and the candidate signal with the highest R^2 value is considered statistically significant if the corresponding p-value is less than or equal to the predefined significance level α (see 21). This TAA signal will then be included in the model and added to the k preselected signals. As a result, the amount of candidate signals, $M-k$, is decreased by one. In such a case, the algorithm starts over from stage 1, and then the remaining candidate signals will try to explain the residuals of this new augmented model. However, if the threshold for statistical significance is not exceeded, the algorithm terminates and concludes that the preselected k variables are the only statistically significant signals for the TAA model.

3.3 Backtesting

TAA strategies are evaluated in terms of risk-adjusted returns generated from out-of-sample **backtests**. Backtests apply historical TAA signal values and generate returns based on trades that would have occurred in the past using the rules defined by the given TAA strategy. Transaction costs and taxes are excluded, i.e., the markets are assumed to be frictionless, and thereby backtests can be conducted by using the stationary panel data (25). As a high measure of a signal indicates either positive or negative future returns, the signals indicating a bearish market are inverted so that high measure in the data always indicates positive market return.

Some constructed test variables can have only two or three values, while others have an unlimited continues scale of different measures. For the integer variables, the backtested trading rule is simple: 1 indicates a long position, -1 a short position and 0 no position. However, in order to ease the interpretation of the other signals, each data point of a continuous variable is transformed into a **rolling percentile rank** (RPR) score. Rolling percentile rank reflects the percentage of previous data points that are equal to or lower than the given value. Rolling percentile rank is thereby an out-of-sample measure because it uses only data that would have been available at any given point of time, i.e., each z_t uses only values $x_1, x_2 ... x_t$.

$$x_t^{RPR} = \frac{|x_i \leq x_t : i \in \{1, 2...t\}|}{|x_1, x_2...x_t|} * 100\% \tag{32}$$

1: Rolling percentile rank

After the RPR transformation, a simple trading rule can also be applied to continuous variables. If a signal's RPR value exceeds 75%, a long position is taken. Correspondingly, a short position is taken if the RPR value is below 25%, and no position is taken otherwise.

Algorithm 1: Backtest trading rule for continuous variables

if $x_t^{RPR} \geq 75\%$ **then**
  $z_t = y_t$
else if $x_t^{RPR} \leq 25\%$ **then**
  $z_t = -y_t$
else
  $z_t = 0$
end

Finally, the risk-adjusted returns are measured with the **sharpe ratio** (SR)

$$SR = \frac{R - R_f}{\sigma}, \tag{33}$$

which is the average annualized return (R) above the risk-free rate (R_f) divided by returns' annual standard deviation, i.e., volatility. The risk-free rate for the Sharpe ratios is approximated with the 1-Month London Interbank Offered Rate (LIBOR).

3.4 Test setting

Variable construction is presented in the Table 4. Each TAA signal is expected to have either a positive or negative correlation with the future market return. The expected correlations are based on theories and rationales proposed in the previous literature. Each TAA signal is associated with a quantitative test variable which replicates the test variables from previously published papers. However, some data transformations are required in order to obtain stationarity and the desired final forms for the test variables (see section 6 for regression diagnostics).

Data is presented in the table 5. Various sources are required to construct the appropriate variables for the US and European markets. Primary sources include the financial databases of **Bloomberg L.P.** and **Thomson Reuters Corporation**. However, some economic data has also been retrieved from the researchers' web sites and central banks' databases directly. **Federal Reserve Economic Data** (FRED) refers to the database maintained by the research division of the Federal Reserve Bank of St. Louis, and **ECB Statistical Data Warehouse** (SDW) refers to the ECB's online data delivery service. In general, the central bank data is preferred because it is assumed to be most reliable and accurate. However, some ECB's data is only available for the euro area (EA) and not for the whole of Europe.

4 Results and discussion

The correlation matrices are presented in the Tables 6 and 7. In multiple hypothesis testing, correlations represent the statistical dependency structure of the multiple tests. Correlations show that most TAA signals exhibit strong positive or negative dependency. Therefore, neither complete dependence nor complete independence between the tests would be an appropriate assumption for the TAA model creation. Testing each signal, for example, at 5% significance level, would have substantially higher than 5% probability for one or more false discoveries. In turn, controlling tests, for example, with a classical Bonferroni correction ($p_i \leq \frac{\alpha}{m}$), would require p-value practically zero, which is an overly conservative requirement for signals that are not entirely independent either.

Whether causal or not (see section 1.2), correlations also reveal interesting statistical relationships between different signal pairs. In line with previous findings of Asness and Ilmanen (2017), technical indicators, i.e., the Dow theory and momentum, are negatively correlated with valuation ratios (assuming price in the denominator) and the Fed model. The dividend yield and the payout ratio are, in turn, surprisingly weakly related. It seems that the pair behaves somewhat differently after all. Correlations among the real activity indicators are mostly consistent with the proposed economic theories (see section 2.3). For example, the GDP gap and inflation have a positive correlation, which supports the rationale that an overheating economy increases the price level. Inflation and the yield curve have a negative correlation, which, in turn, supports the explanation that the central bank' monetary policy and inflation are interlinked. The balance sheet of the central bank, however, is positively correlated with excess reserves and M1 monetary aggregate but, the relationship with the yield curve is surprisingly weak. The sign of the correlation is positive in the US and negative in the EA. Other signals, such as sentiment indicators and seasonal indicators, have mostly low correlations.

TAA model creation is presented in the Tables 8-13. The left-sided tables show the variable selection process, and the right-sided tables show the created TAA model. The tests were made for both US and European markets with three different prediction horizons: week, month and quarter. The time periods were determined by the shortest time series in the data. However, in order to guarantee a sufficient warm-up period, rolling z-score values used also data before the beginning of the presented time period.

In order to illustrate the effect of multiple testing, each signal is associated with a p-value which would have corresponded the signal's R^2 value in an individual test. The p-values are obtained from the **Student's t-distribution**, which is a widely used distribution in statistical analysis. The t-distribution assumes that the error term is normally distributed, and therefore the t-distribution is also symmetric and bell-shaped. However, the t-distribution has heavier tails than the normal distribution because t-distribution takes the sample size into account. In general, the smaller the sample, the fatter the tails. T-distribution becomes eventually identical to the normal distribution when the sample size approaches infinity.

Several signals could be declared statistically significant by ignoring the fact that multiple signals and samples were tested. However, the last line of the left-sided tables show the threshold for statistical significance when the multiple testing is taken into account. The central bank balance sheet is then the only signal which exceeds the multiple testing hurdle and could be thereby declared statistically significant. However, the balance sheet signal seems to be significant only in the US and only with a week prediction horizon. Therefore, only one TAA model is created. The sign of the coefficient is, however, negative, which contradicts the expected correlation with future returns (see Table 4). OMO and QE activities were assumed to raise prices, and thereby predict bullish market (see section 2.3.4). The finding is, nevertheless, exciting and could be explained, for example, with the signalling effect channel. OMO and QE announcement may signal that expansionary monetary policies are required because the central bank is uncertain about the future economic conditions.

However, even though the other signals are not statistically significant, some of the R^2 values are still rather high. Depending on the sample, the shadow banking sector, ETF flow and the M1 monetary aggregate seem to explain a substantial amount of the returns' variance. However, as multiple tests were made, it is not surprising that some statistical predictability starts to emerge.

TAA strategy backtests are presented in the Tables 14-19. The trade signal, i.e., whether the signal is backtested as a "bullish" or "bearish" indicator, is determined by the expected correlation with future returns (see Table 4). The actual correlations, however, show that with several signals it actually would have been more profitable to trade just the opposite direction than prior literature suggested. However, few signals, such as the inflation rate and ETF flow, seem to beat the passive buy and hold strategy occasionally. Though the transaction costs were excluded, and the index performance is rather low due to the dot-com crash at the beginning of the millennium.

5 Conclusion

Tactical asset allocation (TAA) is a dynamic strategy that actively adjusts the portfolio's strategic allocation mix. The objective of TAA is to exploit short-term variation in the expected returns of different asset or sub-asset classes. Market timing is an extreme form of TAA that involves frequent shifts into and out of asset classes in an attempt to time the market peaks and troughs. In systematic TAA strategies, allocation decisions are based on a quantitative forecasting model. In a TAA model, asset class performances are predicted with TAA signals, which are most often different financial and economic variables.

Investors have tried to forecast the equity risk premium as long as the markets have existed. The modern technical analysis began already at the end of 19^{th} century, when Charles Dow developed principles for predicting equity returns with the Dow Jones Industrial Average and the Dow Jones Transportation Average indexes. Since the Dow theory, investors have tried to use also other market value indicators, such as moving averages, momentum, valuation ratios and the Fed model. Investors soon, however, noticed the fundamental difficulty with all TAA and market timing strategies: "too early equals wrong". Since 1960s, researchers have reported promising results also with company fundamentals, such as dividends, payout ratio and earnings. Cognitive theories, such as the prospect theory, anchoring and confirmation bias seemed to explain the market reactions to dividend and earnings announcements, thereby challenging the efficient market hypothesis.

By the 1980s, real activity indicators had started to attract researchers attention. Productivity measures were soon found fragile, but rising inflation and interest rates appeared to predict low stock market returns. The main explanations included the proxy hypothesis and the central bank's conventional monetary policies. An inverted yield curve seemed to predict market crashes because contractionary monetary policy and low inflation expectations were early signs of a forthcoming recession. After millennium, researchers found that unconventional policy measures, in turn, seemed to inflate asset prices. The relationship between quantitative easing and the stock market was not only explained with the credit channel and portfolio rebalancing but also with signalling effects. Some practitioners also explained the economic downturns with theories of debt, credit and leverage cycles, and tried to forecast the business cycle turns by monitoring the prevailing credit conditions.

In the 2010s, investors started to pay attention also to the global economic indicators. The stock markets in export-driven economies seemed to benefit from currency deprecation. Increases in the Baltic Dry Index and metal prices also appeared to predict bull markets through their link to the global industrial activity. Certain currencies and the gold price, in turn, gained a reputation as investors' safe haven. In recent years, investor sentiment indicators, such as sentiment surveys, the VIX, the put–call ratio and ETF flows, have also got researchers interest. Most sentiment indicators, however, have signalled market overreaction and herd-behaviour, and thereby encouraged to trade against the prevailing sentiment. In recent years, researchers have also hinted that the seasonal return patterns, such as the turn of the month effect, could be profitably exploited.

Extensive research on equity market prediction has now been reviewed and retested on post-millennium data. While introducing the test methodology, the basics of statistical hypothesis testing are reviewed. Especially, the concept of type I error, i.e., the probability of a false discovery, is revisited. Due to definition of type I error, the significance level of a single hypothesis test does not equal to the significance level of multiple tests unless the set of tests are completely identical. Thereby, if the significance level in multiple testing is not adjusted appropriately, considering the number of tests performed and their mutual dependency, the test results become seriously spurious and misleading.

The test results imply that false discoveries are endemic in the vast TAA and market timing literature. Multiple testing, sample selection and several other questionable assumptions have established the seemingly promising results of prior literature. The strategies that happened to work well in the backtests of the most cited journal articles, now fail miserably in out-of-sample testing. Few signals, such as the inflation rate and ETF flow, beat occasionally the passive buy and hold strategy, but the results are still not statistically significant if the multiple testing is taken into account. The only statistically significant signal is the the balance sheet of the Federal reserve (FED). However, in contradiction to previous research, the changes in the FED's balance sheet have predicted negative stock market returns.

To conclude, even if the expected equity risk premium might vary over time, investors should not tactically change the equity weight in the strategic investment portfolio. Robust empirical research shows that even in the absence of transaction costs, the systematic TAA and market timing strategies have not been able to generate significant excess returns on out-of-sample periods. However, if investors are still tempted to try such strategies, the inflation rate and ETF fund flows seem to be the most relevant measures to follow in the US and European markets. Further investigation on tactical timing in the emerging markets is left for future research.

References

Adrian, T., Moench, E., & Shin, H. S. (2010). Financial intermediation, asset prices and macroeconomic dynamics. *FRB of New York Staff Report*(422).

Adrian, T., & Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation*, *19*(3), 418–437.

Ahmad, W., & Sharma, S. K. (2018). Testing output gap and economic uncertainty as an explicator of stock market returns. *Research in International Business and Finance*, *45*, 293–306.

Antoniou, A., & Koutmos, G. (2008). Momentum trading: Evidence from futures markets. *Working paper, University of Durham.*

Apergis, N., & Payne, J. (2013). New evidence on the information and predictive content of the Baltic Dry Index. *International Journal of Financial Studies*, *1*(3), 62–80.

Arnott, R. D., & Bernstein, P. L. (2002). What risk premium is "normal"? *Financial Analysts Journal*, *58*(2), 64–85.

Asness, C. (2002). Fight the fed model: the relationship between stock market yields, bond market yields, and future returns. *Bond Market Yields, and Future Returns (December 2002).*

Asness, C., & Ilmanen, A. (2017). Market timing: Sin a little resolving the valuation timing puzzle. *Journal of Investment Management*, *15*(3), 23–40.

Baek, C. (2019). How are gold returns related to stock or bond returns in the us market? evidence from the past 10-year gold market. *Applied Economics*, 1–8.

Bakshi, G., Panayotov, G., & Skoulakis, G. (2010). The Baltic Dry Index as a predictor of global stock returns, commodity returns, and global economic activity. *Commodity Returns, and Global Economic Activity (October 1, 2010).*

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 159–178.

Bekaert, G., & Wang, X. (2010). Inflation risk and the inflation risk premium. *Economic Policy*, *25*(64), 755–806.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Benjamini, Y., Yekutieli, D., et al. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*(4), 1165–1188.

Bonferroni, C. E., Bonferroni, C., & Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita'. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze.*

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business*, *78*(2), 405–440.

Brunnermeier, M. K., & Pedersen, L. H. (2008). Market liquidity and funding liquidity. *Review of Financial Studies*, *22*(6), 2201–2238.

Campbell, J. Y., & Vuolteenaho, T. (2004). Inflation illusion and stock prices. *American Economic Review*, *94*(2), 19–23.

Chava, S., Park, H., & Gallmeyer, M. (2010). *Credit conditions and expected stock returns* (Tech. Rep.). Citeseer.

Cooper, I., & Priestley, R. (2008). Time-varying risk premiums and the output gap. *Review of Financial Studies*, *22*(7), 2801–2833.

Dalio, R. (2012). How the economic machine works. *Economic Principles.*

Dalio, R. (2018). *Principles for navigating big debt crises: The archetypal big debt cycle.* Bridgewater.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *Journal of Finance*, *53*(6), 1839–1885.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, *50*(272), 1096–1121.

Erb, C. B., & Harvey, C. R. (2013). The golden dilemma. *Financial Analysts Journal*, *69*(4), 10–42.

Etula, E., Rinne, K., Suominen, M., & Vaittinen, L. (2015). Dash for cash: Month-end liquidity needs and the predictability of stock returns. *Finance Down Under 2016 Building on the Best from the Cellars of Finance.*

Evensky, H., Horan, S. M., & Robinson, T. R. (2011). *The new wealth management: The financial advisor's guide to managing and investing client assets* (Vol. 28). John Wiley & Sons.

Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *American Economic Review*, *71*(4), 545–565.

Fang, J., Qin, Y., & Jacobsen, B. (2014). Technical market indicators: An overview. *Journal of Behavioral and Experimental Finance*, *4*, 25–56.

Fisher, I. (1930). *Theory of interest: As determined by impatience to spend income and opportunity to invest it.* Augustusm Kelly Publishers, Clifton.

Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, *88*(2), 299–322.

Gambetti, L., & Musso, A. (2017). The macroeconomic impact of the ecb's expanded asset purchase programme (app).

Geanakoplos, J. (2010). The leverage cycle. *NBER macroeconomics annual*, *24*(1), 1–66.

Good, I. J. (1958). Significance tests in parallel and in series. *Journal of the American Statistical Association*, *53*(284), 799–813.

Graham, B., Dodd, D. L. F., Cottle, S., et al. (1934). *Security analysis.* McGraw-Hill New York.

Greenspan, A. (2008). *The age of turbulence: Adventures in a new world.* Penguin.

Harvey, C. R. (1989). Forecasts of economic growth from the bond and stock markets. *Financial Analysts Journal*, *45*(5), 38–45.

Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–802.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Huang, D., & Kilic, M. (2019). Gold, platinum, and expected stock returns. *Journal of Financial Economics*, *132*(3), 50–75.

Ibbotson, R. G. (2010). The importance of asset allocation. *Financial Analysts Journal*, *66*(2), 18–20.

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

Ilmanen, A. (2012). Expected returns on major asset classes. *CFA Institute Research Foundation*, *1*.

Jacobsen, B., Marshall, B. R., & Visaltanachoti, N. (2018). Stock market predictability and industrial metal returns. *Management Science*.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*, 278.

Lamont, O. (1998). Earnings and expected returns. *Journal of Finance*, *53*(5), 1563–1587.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, *56*(3), 815–849.

Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review*, *46*(2), 97–113.

Mamaysky, H. (2014). The time horizon of price responses to quantitative easing. *Available at SSRN*, *2278567*.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, *7*(1), 77-91.

Martikainen, T., & Puttonen, V. (1996). Call-put signal predicts finnish stock returns. *Applied Economics Letters*, *3*(10), 645–648.

Nagel, S. (2012). Evaporating liquidity. *Review of Financial Studies*, *25*(7), 2005–2039.

Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282.

Schannep, J. (2008). Dow theory for the 21st century. *Technical Indicators for Improving Your Investment Results*, *4*.

Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika*, *40*(1-2), 87–110.

Schmidt, B., & Clayton, R. (2017). Super bowl indicator and equity markets: Correlation not causation. *Journal of Business Inquiry*, *17*(2), 97–103.

Sekmen, F. (2011). Exchange rate volatility and stock returns for the us. *African Journal of Business Management*, *5*(22), 9659–9664.

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, *18*(2), 7–19.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, *40*(3), 777–790.

Shiller, R. J., Fischer, S., & Friedman, B. M. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*, *1984*(2), 457–510.

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*(318), 626–633.

Soros, G. (2015). *The alchemy of finance.* John Wiley & Sons.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288–302.

Stockton, K. A., & Shtekhman, A. (2010). A primer on tactical asset allocation strategy evaluation. *Vanguard Group Inc., Valley Forge, PA.*

Tabs, T. (2010). Using equity etf flows as a contrary leading indicator. *TrimTabs Research Note.*

Tukey, J. W., et al. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*(2), 99–114.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Vivian, A., & Wohar, M. E. (2013). The output gap and stock returns: Do cyclical fluctuations predict portfolio returns? *International Review of Financial Analysis*, *26*, 40–50.

Westfall, P. H., Young, S. S., et al. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Vol. 279). John Wiley & Sons.

Yau, H.-Y., & Nieh, C.-C. (2006). Interrelationships among stock prices of taiwan and japan and NTD/YEN exchange rate. *Journal of Asian Economics*, *17*(3), 535–552.

6 Appendix

Table 4: Variable construction

Variable name	**Expected correlation with future return**	**Theories and rationales**	**Test variable (x_t)**	**Data transformations (z_t)**
Market return	(Not relevant)	(Dependent variable)	$P_{t+prediction\ horizon}$	Log difference
Dow theory	Positive	Industrial sector requires transportation services	$\begin{cases} 1 & \text{if } (P_t^{DJIA} > SMA_{200}^{DJIA}) \\ & \wedge(P_t^{DJTA} > SMA_{200}^{DJTA}) \\ -1 & \text{if } (P_t^{DJIA} < SMA_{200}^{DJIA}) \\ & \wedge(P_t^{DJTA} < SMA_{200}^{DJTA}) \\ 0 & \text{otherwise} \end{cases}$ $SMA \equiv$ Simple moving average (see Equation 1)	(Not required)
Momentum: 3m	Positive	Portfolio rebalancing, the disposition effect, representativeness heuristic, overconfidence bias, stop-loss orders, margin calls, portfolio insurance strategies	$\log\left(\frac{P_t}{P_{t-3m}}\right)$	(Not required)
Momentum: 12m	Positive	(same as above)	$\log\left(\frac{P_t}{P_{t-12m}}\right)$	(Not required)
Shiller PE	Negative	Intuition, narratives	$\frac{P}{E_{\text{real 10y avg.}}}$	Rolling z-score
Dividend yield	Positive	(same as above)	$\frac{D}{P}$	Rolling z-score
Value composite	Negative	(same as above)	$\frac{P}{E}, \frac{P}{E_{1y,\ forecast}}, \frac{P}{D}, \frac{P}{D_{1y,\ forecast}}, \frac{P}{B}, \frac{P}{B_{1y,\ forecast}}$	Composite z-score
Fed model	Positive	Money illusion, reflexivity theory	$\frac{E_{1y,\ forward}}{P} - Y_{10y,\ GOV}$	(Not required)
Payout ratio	Positive	Dividend signalling, prospect theory, Lintner's dividend policy model	$\frac{D}{E}$	Difference
Standardized unexpected earnings	Positive	Anchoring, confirmation bias	$\frac{EPS - EPS_{forecast}}{\sigma(EPS - EPS_{forecast})}$	(Not required)
GDP gap	Negative	Full employment	$\frac{GDP_{actual} - GDP_{potential}}{GDP_{potential}}$	Difference

Consumption to wealth ratio	Positive	Consumers prefer stable consumption	$\frac{c_t}{\beta_a a_t+\beta_y y_t}$	Difference
Inflation target: 2%	Negative	Proxy hypothesis, stagfaltion, deflation, money illusion, accounting and taxation rules	$\pi - 2\%$	(Not required)
Yield curve: 10y-2y	Positive	Smart money, self-fulfilling prophesy	$Y_{10y,\ GOV} - Y_{2y,\ GOV}$	Rolling z-score
Central bank balance sheet	Positive	Real activity, portfolio rebalancing	A_{total}	Log difference
Excess reserves	Positive	(same as above)	R_e	Difference
Monetary aggregate: M1	Positive	Herd-behaviour, availability bias, FOMO	M_1,	Log difference
Debt to GDP ratio	Negative	(same as above)	$\frac{D_{\text{GOV}}+D_{\text{BUS}}+D_{\text{HH}}+D_{\text{NPO}}}{GDP}$	Log difference
Shadow banking	Negative	(same as above)	A_{total}	Difference
Credit impulse	Negative	(same as above)	$\frac{Cr_t-Cr_{t-1}}{GDP_t} - \frac{Cr_{t-1}-Cr_{t-2}}{GDP_{t-1}}$	(Not required)
Credit spread	Negative	(same as above)	$Y_{BAA} - Y_{AAA}$	Rolling z-score
Domestic currency	Negative	Exports and domestic currency negatively related	$\frac{C_{\text{domestic}}}{C_{\text{foreign}}}$	Log difference
Safe haven currency: JPY	Negative	Currency carry, smart money	$\frac{C_{\text{JPY}}}{C_{\text{domestic}}}$	Log difference
Safe haven currency: CHF	Negative	(same as above)	$\frac{C_{\text{CHF}}}{C_{\text{domestic}}}$	Log difference
Baltic Dry Index	Positive	Demand of raw materials drives the world economy	(see Equation 16)	Log difference
Industrial metals	Positive	(same as above)	P_t	Log difference

Gold	Negative	Safe haven asset	P_t	Log difference
Bull-bear ratio	Negative	Overoptimism, herd-behaviour, short-selling limitations	$\frac{\text{Bullish}}{\text{Bullish+Bearish}}$	Difference, Rolling z-score
CBOE Volatility Index	Negative	Market makers and hedgefunds less active	(see Equation 18)	Difference
Put-call ratio	Positive	Excessive sentiment, smart money	$\frac{\text{Put volume}}{\text{Call volume}}$	Rolling z-score
ETF flow	Negative	Dumb money, hedge fund's risk management	$\frac{ETF_{\text{creation}} - ETF_{\text{redemption}}}{\text{Total market cap}}$	Rolling z-score
January effect	Positive	Rebound from December, annual bonus payments	$\begin{cases} 1, \text{ if } t \in \{\text{Jan}\} \\ 0, \text{ otherwise} \end{cases}$	(Not required)
Halloween effect	Positive	Summer holidays	$\begin{cases} 1, \text{ if } t \in \{\text{Nov—Apr}\} \\ 0, \text{ otherwise} \end{cases}$	(Not required)
Turn of month effect	Positive	Pensions, dividends, distributions of mutual funds	$\begin{cases} 1, \text{ if } t \in \{\text{T-3—T+3}\} \\ 0, \text{ otherwise} \end{cases}$ $T \equiv$ last day of the month	(Not required)

Table 5: Data

Variable	Raw data	Frequency	Period	Source
Market return	MSCI US index	Daily	1969-2019	Bloomberg
	MSCI Europe index	Daily	1999-2019	Bloomberg
Dow theory	DJIA Index	Daily	1917-2019	Thomson Reuters
	DJTA index	Daily	1988-2019	Thomson Reuters
Momentum: 3m	(see market return)			
Momentum: 12m	(see market return)			
Shiller PE	Cyclically-adjusted price-to-earnings ratio of S&P 500 index	Monthly	1871-2019	Robert Shiller website
Dividend yield	MSCI US dividend yield	Daily	1995-2019	Bloomberg
	MSCI Europe dividend yield	Daily	1999-2019	Bloomberg
Fed model	MSCI US price-earnings ratio	Daily	1995-2019	Bloomberg
	10 year US treasury rate	Daily	1990-2019	Thomson Reuters
Value composite	MSCI US price-earnings ratio	Daily	1995-2019	Bloomberg
	MSCI US price-earnings ratio 12 month forecast	Daily	1995-2019	Bloomberg
	MSCI US dividend yield	Daily	1995-2019	Bloomberg
	MSCI US dividend yield 12 month forecast	Daily	1995-2019	Bloomberg
	MSCI US price-to-book ratio	Daily	1995-2019	Bloomberg
	MSCI US price-to-book ratio 12 month forecast	Daily	1995-2019	Bloomberg
	MSCI Europe price-earnings ratio	Daily	1999-2019	Bloomberg
	MSCI Europe price-earnings ratio 12 month forecast	Daily	1999-2019	Bloomberg
	MSCI Europe dividend yield	Daily	1999-2019	Bloomberg
	MSCI Europe dividend yield 12 month forecast	Daily	1999-2019	Bloomberg
	MSCI Europe price-to-book ratio	Daily	1999-2019	Bloomberg
	MSCI Europe price-to-book ratio 12 month forecast	Daily	1999-2019	Bloomberg
Payout ratio	MSCI US dividend payout ratio	Daily	1995-2019	Bloomberg
	MSCI Europe dividend payout ratio	Daily	1999-2019	Bloomberg
Standardized unexpected earnings	S&P500 Earnings per share - actual surprise: difference between the actual EPS and the average of last broker estimates	Quarterly	1996-2019	Thomson Reuters

GDP gap	Congressional Budget Office (CBO) estimates of GDP gap in the United States	Quarterly	1947-2019	Bloomberg
	International monetary fund (IMF) estimates of GDP gap in Europe	Quarterly	1992-2020	Bloomberg
Consumption to wealth ratio	(See "Monetary Policy and Asset Valuation" by Bianchi, Lettau, and Ludvigson)	Quarterly	1951-2019	Martin Lettau website
Inflation target: 2%	Annual percent change in consumer price index (CPI) in the US	Monthly	1948-2019	FRED
	Annual percent change in CPI in the EA	Monthly	1996-2019	FRED
Yield curve: 10y-2y	10 year US treasury rate	Daily	1977-2019	Bloomberg
	2 year US treasury rate	Daily	1977-2019	Bloomberg
	10 year Germany governement bond yield	Daily	1990-2019	Bloomberg
	2 year Germany governement bond yield	Daily	1990-2019	Bloomberg
Central bank balance sheet	Total assets of Federal Reserve (FED) balance sheet	Weekly	2002-2019	FRED
	Total assets of Federal Reserve (FED) balance sheet	Weekly	1994-2019	Bloomberg
	Total assets of European Central Bank (ECB) balance sheet	Monthly	1999-2019	ECB SDW
Excess reserves	Excess liquidity of depository institutions in the US	Weekly	1984-2019	Bloomberg
	Excess liquidity of depository institutions in the EA	Weekly	1999-2019	Bloomberg
Monetary aggregate: M1	Sum of currency in circulation and overnight deposits in the US	Monthly	1981-2019	Bloomberg
	Sum of currency in circulation and overnight deposits in the EA	Monthly	1980-2019	Bloomberg
Debt to GDP ratio	Total loans of all sectors in the US	Quarterly	1945-2019	FRED
	Total debt securities of all sectors in the US	Quarterly	1945-2019	FRED
Shadow banking	Total value of shadow banking sector in the US	Quarterly	1999-2019	Bloomberg
	Total value of shadow banking sector in the EA	Quarterly	2003-2019	Bloomberg
Credit impulse	New credit issued in the US	Quarterly	1949-2019	Bloomberg
	New credit issued in the EA	Quarterly	2000-2019	Bloomberg
Credit spread	Aaa Corporate Bond Yield in the US	Daily	1983-2019	Bloomberg
	Baa Corporate Bond Yield in the US	Daily	1986-2019	Bloomberg
	Aaa Corporate Bond Yield in the EA	Daily	1994-2019	Bloomberg

	Baa Corporate Bond Yield in the EA	Daily	1994-2019	Bloomberg
Domestic currency	The nominal effective exchange rate of US dollar	Daily	1994-2019	Bloomberg
	The nominal effective exchange rate of the euro	Daily	1996-2019	Bloomberg
Safe haven currency: JPY	USD/JPY nominal exchange rate	Daily	1971-2019	Bloomberg
	EUR/JPY nominal exchange rate	Daily	1986-2019	Bloomberg
Safe haven currency: CHF	USD/CHF nominal exchange rate	Daily	1971-2019	Bloomberg
	EUR/CHF nominal exchange rate	Daily	1999-2019	Bloomberg
Baltic Dry Index	Baltic Dry Index	Daily	1985-2019	Bloomberg
Industrial metals	Bloomberg Industrial Metals Subindex	Daily	1991-2019	Bloomberg
Gold	The spot gold price in US dollars	Daily	1920-2019	Bloomberg
Bull-bear ratio	American Association of Individual Investors (AAII) Investor Sentiment Survey	Weekly	1987-2019	Bloomberg
CBOE Volatility Index	CBOE Volatility Index	Daily	1990-2019	Bloomberg
Put-call ratio	Volume put-call ratio of equity and index options traded on CBOE	Daily	1995-2019	Thomson Reuters
ETF flow	Net of created and redeemed new shares of ETFs listed in the US, trailing quartile	Weekly	1996-2019	Bloomberg
	MSCI US index	Daily	1969-2019	Bloomberg
	Net of created and redeemed new shares of ETFs listed in the EA, trailing quartile	Weekly	2000-2019	Bloomberg
	MSCI Europe index	Daily	1999-2019	Bloomberg
January effect	(see market return)			
Halloween effect	(see market return)			
Turn of month effect	(see market return)			

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	Dow theory	1	0.66	0.44	0.27	-0.43	0.42	-0.21	-0.07	0.07	0.24	0.03	-0.06	0.04	-0.07	-0.01	-0.03	0.06	-0.26	0.25	-0.46	-0.11	0.1	-0.04	0.08	0.18	0	0.04	-0.1	-0.39	-0.09	0.01	0.08	0.03
2	Momentum: 3m	0.66	1	0.46	0.27	-0.53	0.49	-0.13	-0.12	0.07	0.21	0.04	-0.15	0.03	-0.16	-0.07	-0.03	0.08	-0.36	0.21	-0.55	-0.12	0.13	-0.03	0.14	0.2	0.01	0.01	-0.08	-0.34	-0.33	0.01	-0.01	0
3	Momentum: 12m	0.44	0.46	1	0.53	-0.87	0.59	-0.18	-0.09	0.51	0.66	0.11	0.23	-0.05	-0.13	-0.05	-0.03	0.21	-0.15	0.6	-0.83	-0.04	0.08	0	0.06	0.09	-0.01	0	-0.02	-0.15	-0.2	-0.01	0	0.01
4	Shiller PE	0.27	0.27	0.53	1	-0.71	0.88	-0.44	-0.03	0.53	0.43	0.06	0.13	-0.6	-0.16	-0.11	-0.05	0.31	0.07	0.38	-0.65	-0.01	0.04	0.01	0.04	0.03	-0.03	-0.01	0.01	-0.08	-0.08	0	0.01	0
5	Dividend yield	-0.43	-0.53	-0.87	-0.71	1	-0.69	0.18	0.08	-0.62	-0.7	-0.13	-0.27	0.22	0.16	0.09	0.03	-0.36	0.07	-0.69	0.88	0.07	-0.08	0.01	-0.04	-0.1	0	-0.01	0.04	0.13	0.27	0.02	0.06	-0.01
6	Value composite	0.42	0.49	0.59	0.88	-0.69	1	-0.41	-0.04	0.35	0.38	0.06	-0.16	-0.44	-0.19	-0.1	-0.03	0.07	-0.11	0.31	-0.69	-0.05	0.09	0	0.09	0.09	-0.01	0.01	-0.06	-0.24	-0.18	-0.01	0.01	0
7	Fed model	-0.21	-0.13	-0.18	-0.44	0.18	-0.41	1	0.02	-0.09	-0.08	-0.02	-0.05	0.23	0.05	0	0.01	-0.09	0.11	-0.03	0.28	0.02	-0.05	-0.01	-0.04	-0.07	0.03	0.01	0.01	0.08	-0.08	0.01	0.01	0.02
8	Payout ratio	-0.07	-0.12	-0.09	-0.03	0.08	-0.04	0.02	1	0	-0.1	0.05	-0.03	-0.03	0.08	-0.04	-0.02	0.01	0.06	-0.05	0.09	0.02	0.01	-0.02	-0.05	0	0.01	-0.09	-0.03	-0.03	0.02	0.05	-0.02	0.01
9	Standardized unexpected earnings	0.07	0.07	0.51	0.53	-0.62	0.35	-0.09	0	1	0.44	0.06	0.34	-0.39	-0.13	-0.06	-0.04	0.6	0.22	0.59	-0.53	-0.05	0	-0.04	0.03	0.02	0.03	0	0.02	0.12	-0.02	0.01	0.01	-0.02
10	GDP gap	0.24	0.21	0.66	0.43	-0.7	0.38	-0.08	-0.1	0.44	1	0.15	0.22	-0.07	-0.06	-0.04	0	0.25	-0.08	0.57	-0.6	0	-0.01	-0.01	-0.03	0.02	-0.02	-0.01	0.03	-0.01	-0.05	-0.02	-0.01	0.01
11	Consumption to wealth ratio	0.03	0.04	0.11	0.06	-0.13	0.06	-0.02	0.05	0.06	0.15	1	0.07	-0.01	-0.03	-0.03	-0.02	0.1	-0.03	0.09	-0.21	0.01	0.02	0.02	0.02	0.03	0.03	0	0.03	-0.02	-0.18	0.03	-0.2	-0.01
12	Inflation target: 2%	-0.06	-0.15	0.23	0.13	-0.27	-0.16	-0.05	-0.03	0.34	0.22	0.07	1	-0.24	0.09	-0.01	-0.02	0.41	0.25	0.24	-0.16	0.02	0	0.03	-0.06	-0.03	0.01	0	0.06	0.19	0.07	-0.02	-0.02	0.01
13	Yield curve: 10y-2y	0.04	0.03	-0.05	-0.6	0.22	-0.44	0.23	-0.03	-0.39	-0.07	-0.01	-0.24	1	0.13	0.1	0.06	-0.36	-0.21	-0.11	0.21	-0.01	-0.03	-0.03	-0.03	0	0	0.01	-0.01	-0.11	0.03	0	0.01	0
14	Central bank balance sheet	-0.07	-0.16	-0.13	-0.16	0.16	-0.19	0.05	0.08	-0.13	-0.06	-0.03	0.09	0.13	1	0.22	0.08	-0.07	0.02	-0.08	0.18	0.01	-0.06	-0.04	-0.2	-0.12	-0.01	0.05	0.09	0.03	0.26	-0.08	0.03	-0.01
15	Excess reserves	-0.01	-0.07	-0.05	-0.11	0.09	-0.1	0	-0.04	-0.06	-0.04	-0.03	-0.01	0.1	0.22	1	0.02	-0.05	0.04	-0.02	0.07	-0.02	-0.06	-0.06	-0.12	-0.06	0	-0.03	0.03	-0.02	0.09	0.05	0.03	-0.09
16	Monetary aggregate: M1	-0.03	-0.03	-0.03	-0.05	0.03	-0.03	0.01	-0.02	-0.04	0	-0.02	-0.02	0.06	0.08	0.02	1	-0.05	0	-0.02	0.06	0	0.03	0	-0.03	0.03	0.03	-0.01	-0.01	-0.02	0.07	-0.04	-0.01	-0.06
17	Debt to GDP ratio	0.06	0.08	0.21	0.31	-0.36	0.07	-0.09	0.01	0.6	0.25	0.1	0.41	-0.36	-0.07	-0.05	-0.05	1	0.11	0.2	-0.25	-0.08	0	-0.02	0.03	0.09	0.05	0.01	0	0.07	-0.05	0	0.05	0
18	Shadow banking	-0.26	-0.36	-0.15	0.07	0.07	-0.11	0.11	0.06	0.22	-0.08	-0.03	0.25	-0.21	0.02	0.04	0	0.11	1	0.05	0.17	0.03	-0.05	0	-0.07	-0.11	0.02	-0.02	0.01	0.19	0.28	0.03	0.01	0
19	Credit impulse	0.25	0.21	0.6	0.38	-0.69	0.31	-0.03	-0.05	0.59	0.57	0.09	0.24	-0.11	-0.08	-0.02	-0.02	0.2	0.05	1	-0.62	-0.02	0.02	-0.02	-0.01	0	-0.01	-0.01	0.02	0.03	-0.11	-0.03	-0.02	0.01
20	Credit spread	-0.46	-0.55	-0.83	-0.65	0.88	-0.69	0.28	0.09	-0.53	-0.6	-0.21	-0.16	0.21	0.18	0.07	0.06	-0.25	0.17	-0.62	1	0.03	-0.09	0	-0.08	-0.1	0.01	0.01	-0.02	0.1	0.3	0.04	0.06	0
21	Domestic currency	-0.11	-0.12	-0.04	-0.01	0.07	-0.05	0.02	0.02	-0.05	0	0.01	0.02	-0.01	0.01	-0.02	0	-0.08	0.03	-0.02	0.03	1	0.25	0.6	-0.08	-0.49	-0.49	-0.13	0.39	0.21	0.1	-0.02	-0.01	-0.05
22	Safe haven currency: JPY	0.1	0.13	0.08	0.04	-0.08	0.09	-0.05	0.01	0	-0.01	0.02	0	-0.03	-0.06	-0.06	0.03	0	-0.05	0.02	-0.09	0.25	1	0.37	0.04	0.09	-0.29	0.07	-0.3	-0.19	-0.05	-0.05	0.04	0.02
23	Safe haven currency: CHF	-0.04	-0.03	0	0.01	0.01	0	-0.01	-0.02	-0.04	-0.01	0.02	0.03	-0.03	-0.04	-0.06	0	-0.02	0	-0.02	0	0.6	0.37	1	-0.02	-0.21	-0.41	0.01	0.05	0.03	0.01	0	-0.01	0
24	Baltic Dry Index	0.08	0.14	0.06	0.04	-0.04	0.09	-0.04	-0.05	0.03	-0.03	0.02	-0.06	-0.03	-0.2	-0.12	-0.03	0.03	-0.07	-0.01	-0.08	-0.08	0.04	-0.02	1	0.04	0.06	0.05	-0.1	-0.04	-0.21	-0.18	-0.07	-0.05
25	Industrial metals	0.18	0.2	0.09	0.03	-0.1	0.09	-0.07	0	0.02	0.02	0.03	-0.03	0	-0.12	-0.06	0.03	0.09	-0.11	0	-0.1	-0.49	0.09	-0.21	0.04	1	0.41	0.14	-0.37	-0.26	-0.14	0.02	0.04	0.05
26	Gold	0	0.01	-0.01	-0.03	0	-0.01	0.03	0.01	0.03	-0.02	0.03	0.01	0	-0.01	0	0.03	0.05	0.02	-0.01	0.01	-0.49	-0.29	-0.41	0.06	0.41	1	0.03	-0.09	-0.08	0.01	0.06	0.01	0
27	Bull-bear ratio	0.04	0.01	0	-0.01	-0.01	0.01	0.01	-0.09	0	-0.01	0	0	0.01	0.05	-0.03	-0.01	0.01	-0.02	-0.01	0.01	-0.13	0.07	0.01	0.05	0.14	0.03	1	-0.19	-0.11	-0.01	-0.05	0.01	0.02
28	CBOE Volatility Index	-0.1	-0.08	-0.02	0.01	0.04	-0.06	0.01	-0.03	0.02	0.03	0.03	0.06	-0.01	0.09	0.03	-0.01	0	0.01	0.02	-0.02	0.39	-0.3	0.05	-0.1	-0.37	-0.09	-0.19	1	0.38	0.06	0	-0.03	0.06
29	Put-call ratio	-0.39	-0.34	-0.15	-0.08	0.13	-0.24	0.08	-0.03	0.12	-0.01	-0.02	0.19	-0.11	0.03	-0.02	-0.02	0.07	0.19	0.03	0.1	0.21	-0.19	0.03	-0.04	-0.26	-0.08	-0.11	0.38	1	0.04	-0.05	-0.12	0.02
30	ETF flow	-0.09	-0.33	-0.2	-0.08	0.27	-0.18	-0.08	0.02	-0.02	-0.05	-0.18	0.07	0.03	0.26	0.09	0.07	-0.05	0.28	-0.11	0.3	0.1	-0.05	0.01	-0.21	-0.14	0.01	-0.01	0.06	0.04	1	0.22	0.18	0.01
31	January effect	0.01	0.01	-0.01	0	0.02	-0.01	0.01	0.05	0.01	-0.02	0.03	-0.02	0	-0.08	0.05	-0.04	0	0.03	-0.03	0.04	-0.02	-0.05	0	-0.18	0.02	0.06	-0.05	0	-0.05	0.22	1	0.26	0.01
32	Halloween effect	0.08	-0.01	0	0.01	0.06	0.01	0.01	-0.02	0.01	-0.01	-0.2	-0.02	0.01	0.03	0.03	-0.01	0.05	0.01	-0.02	0.06	-0.01	0.04	-0.01	-0.07	0.04	0.01	0.01	-0.03	-0.12	0.18	0.26	1	-0.01
33	Turn of month effect	0.03	0	0.01	0	-0.01	0	0.02	0.01	-0.02	0.01	-0.01	0.01	0	-0.01	-0.09	-0.06	0	0	0.01	0	-0.05	0.02	0	-0.05	0.05	0	0.02	0.06	0.02	0.01	0.01	-0.01	1

Table 6: Correlation matrix, market: US, time period: 2003-2019

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Momentum: 3m	1	0.49	-0.59	0.52	-0.08	-0.23	-0.2	0.13	-0.3	0.01	0	0.42	0.2	-0.07	0.13	0	0.1	0.2	0.01	-0.06	-0.22	0.02	0.03	0
2	Momentum: 12m	0.49	1	-0.83	0.63	-0.14	-0.29	-0.24	-0.04	-0.19	-0.03	-0.1	0.4	0.02	-0.03	0.09	0.01	0.04	0.11	0	-0.01	-0.09	-0.01	0.02	0
3	Dividend yield	-0.59	-0.83	1	-0.8	0.11	0.18	0.46	0.09	0.23	-0.01	0.11	-0.4	-0.07	0.01	-0.11	-0.02	-0.07	-0.15	-0.01	0.04	0.32	-0.01	-0.04	-0.01
4	Value composite	0.52	0.63	-0.8	1	-0.05	0.02	-0.52	-0.36	-0.27	0.01	-0.04	0.35	-0.02	0.01	0.13	0.04	0.1	0.14	0.02	-0.05	-0.13	-0.01	0.04	0
5	Payout ratio	-0.08	-0.14	0.11	-0.05	1	0.1	0.02	-0.05	-0.06	-0.02	0.01	-0.07	-0.06	-0.06	-0.07	-0.03	-0.02	-0.02	0.02	-0.02	0.05	0.06	0	-0.02
6	GDP gap	-0.23	-0.29	0.18	0.02	0.1	1	0.28	-0.58	0.15	0.06	0.03	-0.42	-0.08	0.05	0.01	0.05	-0.03	-0.07	-0.01	0.03	0.06	0.01	0.01	-0.02
7	Inflation target: 2%	-0.2	-0.24	0.46	-0.52	0.02	0.28	1	-0.05	0.17	-0.03	0.01	-0.31	0.16	0.05	0	0.04	-0.02	-0.03	0.03	0	0.26	-0.05	-0.08	0.01
8	Yield curve: 10y-2y	0.13	-0.04	0.09	-0.36	-0.05	-0.58	-0.05	1	-0.1	-0.02	0.1	0.4	0.08	-0.01	0.01	-0.01	0.01	0.05	0.01	-0.04	0.01	0.02	0.03	0.01
9	Central bank balance sheet	-0.3	-0.19	0.23	-0.27	-0.06	0.15	0.17	-0.1	1	0.11	0.07	-0.28	-0.03	-0.07	-0.16	-0.05	-0.15	-0.18	-0.05	0.11	0.04	-0.02	-0.06	0.02
10	Excess reserves	0.01	-0.03	-0.01	0.01	-0.02	0.06	-0.03	-0.02	0.11	1	0.01	-0.06	-0.02	-0.02	-0.05	0.01	0.01	-0.01	0.03	0.07	-0.04	-0.04	0.03	0.16
11	Monetary aggregate: M1	0	-0.1	0.11	-0.04	0.01	0.03	0.01	0.1	0.07	0.01	1	0.04	-0.01	0.04	0.03	0.03	-0.02	-0.01	0.07	0.04	0.09	-0.02	0.05	-0.03
12	Shadow banking	0.42	0.4	-0.4	0.35	-0.07	-0.42	-0.31	0.4	-0.28	-0.06	0.04	1	0.16	-0.01	0.02	0	0.01	0.07	0.01	-0.01	-0.09	-0.07	-0.02	0.01
13	Credit spread	0.2	0.02	-0.07	-0.02	-0.06	-0.08	0.16	0.08	-0.03	-0.02	-0.01	0.16	1	0.01	0.01	0.04	0.02	0.05	0	-0.01	-0.04	-0.05	-0.12	-0.01
14	Domestic currency	-0.07	-0.03	0.01	0.01	-0.06	0.05	0.05	-0.01	-0.07	-0.02	0.04	-0.01	0.01	1	0.46	0.17	0.04	0.12	0.29	-0.04	0.08	-0.04	-0.02	-0.01
15	Safe haven currency: JPY	0.13	0.09	-0.11	0.13	-0.07	0.01	0	0.01	-0.16	-0.05	0.03	0.02	0.01	0.46	1	0.31	0.08	0.35	0.13	-0.43	0.02	-0.06	0.03	0.03
16	Safe haven currency: CHF	0	0.01	-0.02	0.04	-0.03	0.05	0.04	-0.01	-0.05	0.01	0.03	0	0.04	0.17	0.31	1	0.03	0.1	-0.05	-0.19	-0.01	-0.04	-0.01	0.02
17	Baltic Dry Index	0.1	0.04	-0.07	0.1	-0.02	-0.03	-0.02	0.01	-0.15	0.01	-0.02	0.01	0.02	0.04	0.08	0.03	1	0.04	0.06	-0.1	-0.04	-0.18	-0.08	-0.05
18	Industrial metals	0.2	0.11	-0.15	0.14	-0.02	-0.07	-0.03	0.05	-0.18	-0.01	-0.01	0.07	0.05	0.12	0.35	0.1	0.04	1	0.41	-0.37	-0.05	0.02	0.04	0.05
19	Gold	0.01	0	-0.01	0.02	0.02	-0.01	0.03	0.01	-0.05	0.03	0.07	0.01	0	0.29	0.13	-0.05	0.06	0.41	1	-0.09	0.02	0.07	0.01	0
20	CBOE Volatility Index	-0.06	-0.01	0.04	-0.05	-0.02	0.03	0	-0.04	0.11	0.07	0.04	-0.01	-0.01	-0.04	-0.43	-0.19	-0.1	-0.37	-0.09	1	0	0	-0.03	0.06
21	ETF flow	-0.22	-0.09	0.32	-0.13	0.05	0.06	0.26	0.01	0.04	-0.04	0.09	-0.09	-0.04	0.08	0.02	-0.01	-0.04	-0.05	0.02	0	1	0.15	0.21	-0.01
22	January effect	0.02	-0.01	-0.01	-0.01	0.06	0.01	-0.05	0.02	-0.02	-0.04	-0.02	-0.07	-0.05	-0.04	-0.06	-0.04	-0.18	0.02	0.07	0	0.15	1	0.26	0.01
23	Halloween effect	0.03	0.02	-0.04	0.04	0	0.01	-0.08	0.03	-0.06	0.03	0.05	-0.02	-0.12	-0.02	0.03	-0.01	-0.08	0.04	0.01	-0.03	0.21	0.26	1	-0.01
24	Turn of month effect	0	0	-0.01	0	-0.02	-0.02	0.01	0.01	0.02	0.16	-0.03	0.01	-0.01	-0.01	0.03	0.02	-0.05	0.05	0	0.06	-0.01	0.01	-0.01	1

Table 7: Correlation matrix, market: Europe

Table 8: TAA model creation, market: US, prediction horizon: week, time period: 2000-2019, significance level: 5%

	R^2	p-value
Dow theory	0.000	0.754
Momentum: 3m	0.001	0.429
Momentum: 12m	0.000	0.881
Shiller PE	0.000	0.798
Dividend yield	0.004	0.046
Value composite	0.002	0.176
Fed model	0.000	0.639
Payout ratio	0.000	0.514
Standardized unexpected earnings	0.000	0.807
GDP gap	0.001	0.226
Consumption to wealth ratio	0.001	0.378
Inflation target: 2%	0.006	0.012
Yield curve: 10y-2y	0.000	0.778
Central bank balance sheet	**0.049**	**0.000**
Excess reserves	0.002	0.196
Monetary aggregate: M1	0.001	0.326
Debt to GDP ratio	0.000	0.871
Shadow banking	0.001	0.346
Credit impulse	0.001	0.345
Credit spread	0.001	0.234
Domestic currency	0.000	0.826
Safe haven currency: JPY	0.013	0.000
Safe haven currency: CHF	0.001	0.354
Baltic Dry Index	0.002	0.204
Industrial metals	0.003	0.092
Gold	0.000	0.485
Bull-bear ratio	0.000	0.828
CBOE Volatility Index	0.012	0.000
Put-call ratio	0.006	0.012
ETF flow	0.013	0.000
January effect	0.001	0.430
Halloween effect	0.002	0.207
Turn of month effect	0.001	0.330
Multiple testing	**0.049**	**0.000**

	Dependent variable:
	y
‘Central bank balance sheet‘	-0.311^{***}
	(0.043)
Constant	0.001^{*}
	(0.001)
Observations	1,015
R^2	0.049
Adjusted R^2	0.048
Residual Std. Error	0.023 (df = 1013)
F Statistic	51.849^{***} (df = 1; 1013)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 9: TAA model creation, market: Europe, prediction horizon: week, time period: 2003-2019, significance level: 5%

	R^2	p-value
Momentum: 3m	0.000	0.768
Momentum: 12m	0.000	0.979
Dividend yield	0.001	0.389
Value composite	0.000	0.798
Payout ratio	0.001	0.339
GDP gap	0.005	0.048
Inflation target: 2%	0.002	0.238
Yield curve: 10y-2y	0.002	0.176
Central bank balance sheet	0.010	0.004
Excess reserves	0.007	0.016
Monetary aggregate: M1	0.001	0.319
Shadow banking	**0.011**	**0.002**
Credit impulse	0.005	0.039
Credit spread	0.006	0.025
Domestic currency	0.002	0.212
Safe haven currency: JPY	0.005	0.034
Safe haven currency: CHF	0.000	0.686
Baltic Dry Index	0.000	0.851
Industrial metals	0.005	0.036
Gold	0.001	0.300
CBOE Volatility Index	0.001	0.486
ETF flow	0.003	0.090
January effect	0.000	0.768
Halloween effect	0.001	0.296
Turn of month effect	0.001	0.290
Multiple testing	**0.024**	**0.000**

Table 10: TAA model creation, market: US, prediction horizon: month, time period: 2000-2019, significance level: 5%

	R^2	p-value
Dow theory	0.000	0.886
Momentum: 3m	0.003	0.408
Momentum: 12m	0.003	0.441
Shiller PE	0.000	0.824
Dividend yield	0.011	0.115
Value composite	0.004	0.366
Fed model	0.000	0.864
Payout ratio	0.000	0.906
Standardized unexpected earnings	0.001	0.653
GDP gap	0.008	0.166
Consumption to wealth ratio	0.006	0.237
Inflation target: 2%	0.028	0.010
Yield curve: 10y-2y	0.000	0.950
Central bank balance sheet	0.061	0.000
Excess reserves	0.000	0.933
Monetary aggregate: M1	0.000	0.894
Debt to GDP ratio	0.000	0.917
Shadow banking	0.012	0.098
Credit impulse	0.004	0.314
Credit spread	0.016	0.056
Domestic currency	0.003	0.440
Safe haven currency: JPY	0.015	0.065
Safe haven currency: CHF	0.001	0.709
Baltic Dry Index	0.018	0.039
Industrial metals	0.027	0.012
Gold	0.001	0.585
Bull-bear ratio	0.000	0.925
CBOE Volatility Index	0.015	0.064
Put-call ratio	0.025	0.016
ETF flow	**0.073**	**0.000**
January effect	0.000	0.973
Halloween effect	0.012	0.101
Multiple testing	**0.311**	**0.000**

Table 11: TAA model creation, market: Europe, prediction horizon: month, time period: 2003-2019, significance level: 5%

	R^2	p-value
Momentum: 3m	0.010	0.170
Momentum: 12m	0.001	0.687
Dividend yield	0.010	0.161
Value composite	0.004	0.353
Payout ratio	0.002	0.522
GDP gap	0.026	0.024
Inflation target: 2%	0.007	0.248
Yield curve: 10y-2y	0.015	0.093
Central bank balance sheet	**0.104**	**0.000**
Excess reserves	0.031	0.014
Monetary aggregate: M1	0.022	0.037
Shadow banking	0.076	0.000
Credit impulse	0.028	0.020
Credit spread	0.043	0.003
Domestic currency	0.011	0.152
Safe haven currency: JPY	0.028	0.019
Safe haven currency: CHF	0.010	0.159
Baltic Dry Index	0.004	0.370
Industrial metals	0.061	0.000
Gold	0.005	0.313
CBOE Volatility Index	0.053	0.001
ETF flow	0.012	0.124
January effect	0.001	0.596
Halloween effect	0.010	0.155
Multiple testing	**0.112**	**0.000**

Table 12: TAA model creation, market: US, prediction horizon: quarter, time period: 2000-2019, significance level: 5%

	R^2	p-value
Dow theory	0.003	0.626
Momentum: 3m	0.010	0.378
Momentum: 12m	0.000	0.879
Shiller PE	0.014	0.314
Dividend yield	0.044	0.067
Value composite	0.026	0.162
Fed model	0.001	0.817
Payout ratio	0.072	0.018
Standardized unexpected earnings	0.003	0.622
GDP gap	0.020	0.215
Consumption to wealth ratio	0.018	0.239
Inflation target: 2%	0.104	0.004
Yield curve: 10y-2y	0.000	0.915
Central bank balance sheet	**0.158**	**0.000**
Excess reserves	0.017	0.265
Monetary aggregate: M1	0.000	0.905
Debt to GDP ratio	0.000	0.951
Shadow banking	0.058	0.036
Credit impulse	0.012	0.338
Credit spread	0.021	0.205
Domestic currency	0.000	0.971
Safe haven currency: JPY	0.007	0.458
Safe haven currency: CHF	0.009	0.412
Baltic Dry Index	0.079	0.013
Industrial metals	0.041	0.076
Gold	0.000	0.914
Bull-bear ratio	0.003	0.649
CBOE Volatility Index	0.000	0.898
Put-call ratio	0.034	0.111
ETF flow	0.095	0.006
January effect	0.001	0.794
Halloween effect	0.008	0.442
Multiple testing	**0.373**	**0.000**

Table 13: TAA model creation, market: Europe, prediction horizon: quarter, time period: 2003-2019, significance level: 5%

	R^2	p-value
Momentum: 3m	0.024	0.219
Momentum: 12m	0.000	0.864
Dividend yield	0.005	0.567
Value composite	0.001	0.787
Payout ratio	0.003	0.656
GDP gap	0.062	0.045
Inflation target: 2%	0.013	0.362
Yield curve: 10y-2y	0.052	0.067
Central bank balance sheet	0.131	0.003
Excess reserves	0.021	0.244
Monetary aggregate: M1	**0.183**	**0.000**
Shadow banking	0.181	0.000
Credit impulse	0.078	0.024
Credit spread	0.065	0.040
Domestic currency	0.002	0.710
Safe haven currency: JPY	0.022	0.236
Safe haven currency: CHF	0.063	0.043
Baltic Dry Index	0.140	0.002
Industrial metals	0.086	0.018
Gold	0.000	0.938
CBOE Volatility Index	0.001	0.841
ETF flow	0.059	0.052
January effect	0.004	0.605
Halloween effect	0.000	0.996
Multiple testing	**0.235**	**0.000**

Table 14: TAA strategy backtests, market: US, prediction horizon: week, time period: 2000-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Dow theory	Bull	-0.01	4.8	0.2	-0.09
Momentum: 3m	Bull	-0.02	14.0	0.7	-0.06
Momentum: 12m	Bull	0.00	48.8	2.1	0.03
Shiller PE	Bear	-0.01	-58.4	-4.4	-0.39
Dividend yield	Bull	0.06	77.0	3.0	0.09
Value composite	Bear	-0.04	-24.3	-1.4	-0.21
Fed model	Bull	0.01	10.6	0.5	-0.07
Payout ratio	Bull	-0.02	4.0	0.2	-0.10
Standardized unexpected earnings	Bull	0.01	0.7	0.0	-0.12
GDP gap	Bear	0.04	-52.3	-3.7	-0.43
Consumption to wealth ratio	Bull	-0.03	-40.9	-2.7	-0.34
Inflation target: 2%	Bear	-0.08	107.9	3.8	0.15
Yield curve: 10y-2y	Bull	-0.01	-57.3	-4.3	-0.49
Central bank balance sheet	Bull	-0.22	-56.1	-4.2	-0.49
Excess reserves	Bull	-0.04	-79.3	-7.8	-0.56
Monetary aggregate: M1	Bull	-0.03	-36.9	-2.3	-0.29
Debt to GDP ratio	Bear	0.01	-34.0	-2.1	-0.27
Shadow banking	Bear	-0.03	78.3	3.0	0.11
Credit impulse	Bear	0.03	-36.4	-2.3	-0.29
Credit spread	Bear	-0.04	137.7	4.6	0.20
Domestic currency	Bear	-0.01	23.1	1.1	-0.04
Safe haven currency: JPY	Bear	-0.12	221.8	6.2	0.36
Safe haven currency: CHF	Bear	-0.03	4.3	0.2	-0.11
Baltic Dry Index	Bull	0.04	91.1	3.4	0.12
Industrial metals	Bull	-0.05	-74.4	-6.8	-0.60
Gold	Bear	-0.02	71.6	2.8	0.09
Bull-bear ratio	Bear	-0.01	21.8	1.0	-0.05
CBOE Volatility Index	Bear	0.11	-43.5	-2.9	-0.32
Put-call ratio	Bull	0.08	123.6	4.2	0.21
ETF flow	**Bear**	**-0.12**	**411.9**	**8.8**	**0.54**
January effect	Bull	-0.02	-10.2	-0.6	-0.53
Halloween effect	Bull	0.04	146.3	4.7	0.22
Turn of month effect	Bull	-0.03	-13.4	-0.7	-0.30
Buy and hold			**106.4**	**3.8**	**0.13**

Table 15: TAA strategy backtests, market: Europe, prediction horizon: week, time period: 2003-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Momentum: 3m	Bull	-0.01	-0.6	0.0	-0.14
Momentum: 12m	Bull	0.00	18.2	1.0	-0.06
Dividend yield	Bull	-0.03	-30.1	-2.2	-0.31
Value composite	Bear	0.01	-20.1	-1.4	-0.25
Payout ratio	Bull	-0.03	101.3	4.4	0.22
GDP gap	Bear	-0.07	91.4	4.1	0.26
Inflation target: 2%	**Bear**	**-0.04**	**246.6**	**7.9**	**0.45**
Yield curve: 10y-2y	Bull	0.05	-22.9	-1.6	-0.31
Central bank balance sheet	Bull	-0.10	-62.9	-5.9	-0.71
Excess reserves	Bull	-0.08	-40.5	-3.1	-0.35
Monetary aggregate: M1	Bull	-0.03	-12.2	-0.8	-0.22
Shadow banking	Bear	0.10	-51.9	-4.4	-0.52
Credit impulse	Bear	-0.07	116.5	4.9	0.22
Credit spread	Bear	0.08	-6.9	-0.4	-0.27
Domestic currency	Bear	0.04	-23.8	-1.7	-0.29
Safe haven currency: JPY	Bear	-0.07	30.7	1.7	-0.01
Safe haven currency: CHF	Bear	-0.01	-19.8	-1.3	-0.24
Baltic Dry Index	Bull	0.01	-23.7	-1.6	-0.27
Industrial metals	Bull	-0.07	-67.7	-6.7	-0.67
Gold	Bear	-0.04	34.1	1.8	0.00
CBOE Volatility Index	Bear	0.02	-13.8	-0.9	-0.20
ETF flow	Bear	-0.06	88.4	4.0	0.16
January effect	Bull	-0.01	0.2	0.0	-0.40
Halloween effect	Bull	0.04	105.7	4.5	0.22
Turn of month effect	Bull	-0.04	-14.2	-0.9	-0.36
Buy and hold			**94.3**	**4.2**	**0.14**

Table 16: TAA strategy backtests, market: US, prediction horizon: month, time period: 2000-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Dow theory	Bull	-0.01	16.6	0.8	-0.12
Momentum: 3m	Bull	0.05	40.7	1.8	-0.05
Momentum: 12m	Bull	0.05	42.0	1.8	-0.05
Shiller PE	Bear	0.01	-69.7	-6.0	-0.62
Dividend yield	Bull	0.10	68.3	2.7	0.02
Value composite	Bear	-0.06	6.4	0.3	-0.16
Fed model	Bull	-0.01	25.0	1.2	-0.09
Payout ratio	Bull	-0.01	-24.5	-1.4	-0.31
Standardized unexpected earnings	Bull	-0.03	-6.6	-0.4	-0.24
GDP gap	Bear	0.09	-48.4	-3.3	-0.53
Consumption to wealth ratio	Bull	-0.08	-53.0	-3.8	-0.54
Inflation target: 2%	Bear	-0.17	76.1	3.0	0.04
Yield curve: 10y-2y	Bull	0.00	-57.9	-4.3	-0.64
Central bank balance sheet	Bull	-0.25	-6.0	-0.3	-0.23
Excess reserves	Bull	0.01	-6.5	-0.3	-0.20
Monetary aggregate: M1	Bull	-0.01	49.6	2.1	-0.03
Debt to GDP ratio	Bear	0.01	-22.5	-1.3	-0.31
Shadow banking	Bear	-0.11	113.8	4.0	0.14
Credit impulse	Bear	0.07	-35.9	-2.3	-0.41
Credit spread	Bear	-0.12	96.3	3.5	0.09
Domestic currency	Bear	-0.05	16.9	0.8	-0.13
Safe haven currency: JPY	Bear	0.12	-53.7	-3.9	-0.58
Safe haven currency: CHF	Bear	-0.02	40.3	1.8	-0.06
Baltic Dry Index	Bull	0.14	62.6	2.5	0.01
Industrial metals	Bull	0.16	116.1	4.0	0.12
Gold	Bear	-0.04	121.0	4.2	0.13
Bull-bear ratio	Bear	0.01	7.5	0.4	-0.23
CBOE Volatility Index	Bear	-0.12	-43.9	-2.9	-0.43
Put-call ratio	Bull	0.16	104.1	3.7	0.14
ETF flow	Bear	-0.27	-70.9	-6.2	-0.82
January effect	Bull	0.00	6.6	0.3	-0.48
Halloween effect	**Bull**	**0.11**	**169.8**	**5.2**	**0.23**
Buy and hold			**109.7**	**3.9**	**0.09**

Table 17: TAA strategy backtests, market: Europe, prediction horizon: month, time period: 2003-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Momentum: 3m	Bull	0.10	49.2	2.5	0.00
Momentum: 12m	Bull	0.03	21.5	1.2	-0.11
Dividend yield	Bull	-0.10	-33.9	-2.5	-0.43
Value composite	Bear	0.07	-47.3	-3.9	-0.54
Payout ratio	Bull	0.05	57.6	2.8	0.04
GDP gap	Bear	-0.16	79.2	3.7	0.17
Inflation target: 2%	**Bear**	**-0.08**	**204.1**	**7.1**	**0.44**
Yield curve: 10y-2y	Bull	0.12	10.2	0.6	-0.19
Central bank balance sheet	Bull	-0.32	-65.9	-6.4	-0.98
Excess reserves	Bull	-0.18	-66.8	-6.6	-0.80
Monetary aggregate: M1	Bull	-0.15	-0.4	0.0	-0.22
Shadow banking	Bear	0.27	-63.0	-5.9	-0.87
Credit impulse	Bear	-0.17	76.4	3.5	0.09
Credit spread	Bear	0.21	-10.8	-0.7	-0.42
Domestic currency	Bear	0.10	-54.3	-4.7	-0.71
Safe haven currency: JPY	Bear	0.17	-34.5	-2.6	-0.50
Safe haven currency: CHF	Bear	0.10	-39.5	-3.0	-0.54
Baltic Dry Index	Bull	0.06	-21.6	-1.5	-0.33
Industrial metals	Bull	0.25	214.2	7.3	0.41
Gold	Bear	-0.07	83.2	3.8	0.12
CBOE Volatility Index	Bear	-0.23	104.8	4.5	0.17
ETF flow	Bear	-0.11	55.7	2.8	0.03
January effect	Bull	0.04	15.2	0.9	-0.34
Halloween effect	Bull	0.10	107.8	4.6	0.21
Buy and hold			**88.9**	**4.0**	**0.11**

Table 18: TAA strategy backtests, market: US, prediction horizon: quarter, time period: 2000-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Dow theory	Bull	-0.06	-18.0	-1.0	-0.22
Momentum: 3m	Bull	0.10	38.5	1.7	-0.05
Momentum: 12m	Bull	0.02	81.0	3.2	0.05
Shiller PE	Bear	-0.12	-39.5	-2.6	-0.35
Dividend yield	Bull	0.21	74.8	3.0	0.04
Value composite	Bear	-0.16	-19.3	-1.1	-0.27
Fed model	Bull	0.03	34.5	1.6	-0.06
Payout ratio	Bull	-0.27	-67.8	-5.8	-0.67
Standardized unexpected earnings	Bull	-0.06	2.3	0.1	-0.19
GDP gap	Bear	0.14	-23.9	-1.4	-0.30
Consumption to wealth ratio	Bull	-0.14	-47.4	-3.3	-0.49
Inflation target: 2%	**Bear**	**-0.32**	**251.5**	**6.8**	**0.33**
Yield curve: 10y-2y	Bull	0.01	-52.3	-3.8	-0.52
Central bank balance sheet	Bull	-0.40	-61.1	-4.8	-0.59
Excess reserves	Bull	-0.13	-28.6	-1.8	-0.28
Monetary aggregate: M1	Bull	-0.01	-24.6	-1.5	-0.33
Debt to GDP ratio	Bear	0.01	-20.8	-1.2	-0.29
Shadow banking	Bear	-0.24	7.1	0.4	-0.18
Credit impulse	Bear	0.11	-34.9	-2.2	-0.41
Credit spread	Bear	-0.15	128.1	4.4	0.16
Domestic currency	Bear	0.00	-27.9	-1.7	-0.33
Safe haven currency: JPY	Bear	0.09	-28.5	-1.7	-0.36
Safe haven currency: CHF	Bear	0.09	-49.2	-3.5	-0.50
Baltic Dry Index	Bull	0.28	78.0	3.1	0.05
Industrial metals	Bull	0.20	24.0	1.1	-0.10
Gold	Bear	-0.01	-1.9	-0.1	-0.21
Bull-bear ratio	Bear	0.05	0.9	0.0	-0.22
CBOE Volatility Index	Bear	0.01	-27.3	-1.7	-0.30
Put-call ratio	Bull	0.18	-1.2	-0.1	-0.24
ETF flow	Bear	-0.31	74.4	3.0	0.05
January effect	Bull	0.03	27.4	1.3	-0.17
Halloween effect	Bull	0.09	115.8	4.1	0.13
Buy and hold			**100.5**	**3.7**	**0.08**

Table 19: TAA strategy backtests, market: Europe, prediction horizon: quarter, time period: 2003-2019

	Trade signal	Correlation with future return	Cumulative return (%)	Annualized return (%)	Sharpe ratio
Momentum: 3m	Bull	0.15	87.9	4.0	0.12
Momentum: 12m	Bull	-0.02	51.3	2.6	0.02
Dividend yield	Bull	-0.07	-32.7	-2.4	-0.39
Value composite	Bear	0.03	-40.8	-3.2	-0.45
Payout ratio	Bull	-0.06	-28.6	-2.1	-0.36
GDP gap	Bear	-0.25	78.2	3.7	0.15
Inflation target: 2%	**Bear**	**-0.12**	**166.6**	**6.3**	**0.32**
Yield curve: 10y-2y	Bull	0.23	-6.6	-0.4	-0.30
Central bank balance sheet	Bull	-0.36	-63.3	-6.1	-0.74
Excess reserves	Bull	-0.15	-38.8	-3.0	-0.38
Monetary aggregate: M1	Bull	0.43	197.5	7.0	0.32
Shadow banking	Bear	0.43	-63.2	-6.1	-0.83
Credit impulse	Bear	-0.28	116.8	5.0	0.18
Credit spread	Bear	0.26	-7.5	-0.5	-0.37
Domestic currency	Bear	0.05	72.3	3.5	0.09
Safe haven currency: JPY	Bear	0.15	-40.8	-3.2	-0.45
Safe haven currency: CHF	Bear	0.25	-54.9	-4.9	-0.60
Baltic Dry Index	Bull	0.37	117.1	5.0	0.19
Industrial metals	Bull	0.29	137.7	5.6	0.21
Gold	Bear	0.01	-43.5	-3.5	-0.54
CBOE Volatility Index	Bear	-0.03	-50.3	-4.3	-0.49
ETF flow	Bear	-0.24	83.4	3.9	0.11
January effect	Bull	0.07	37.8	2.0	-0.05
Halloween effect	Bull	0.00	55.2	2.8	0.03
Buy and hold			**85.0**	**3.9**	**0.10**

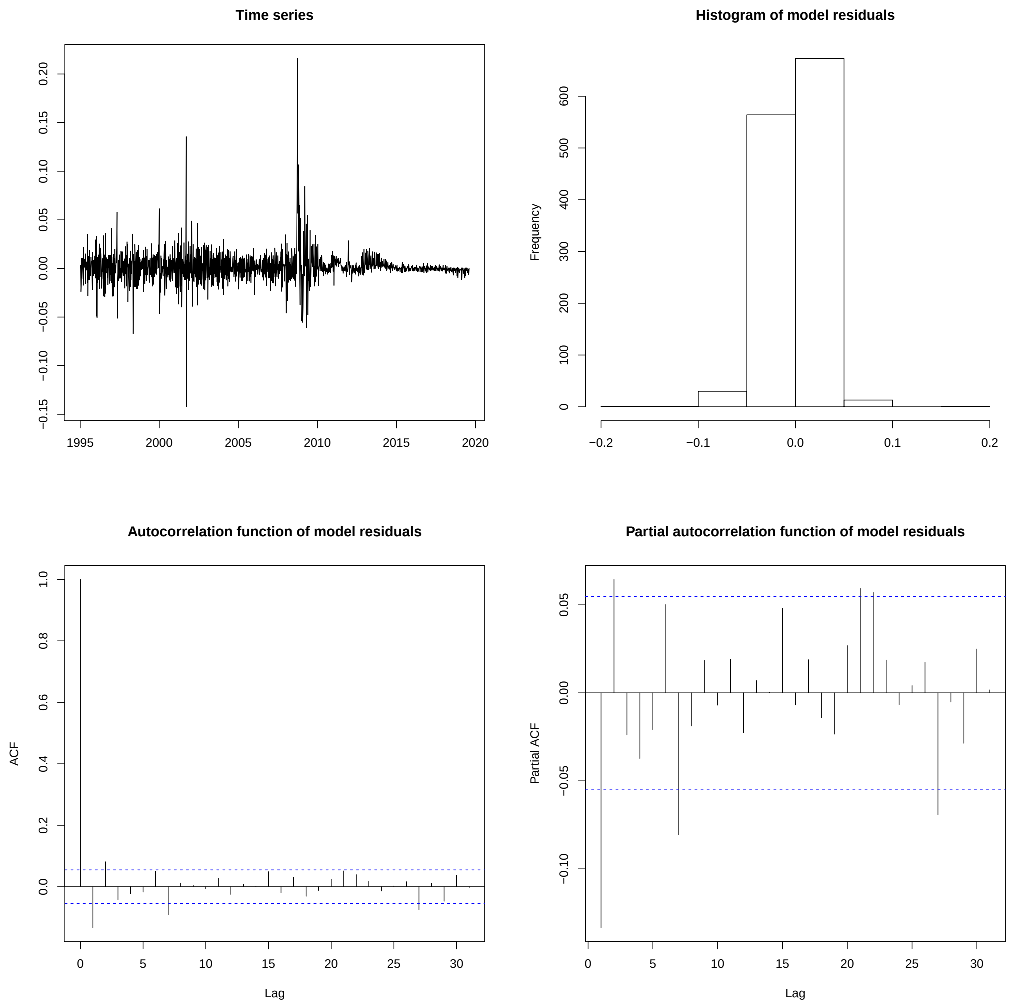

Figure 4: Regression diagnostics: Central bank balance sheet, US

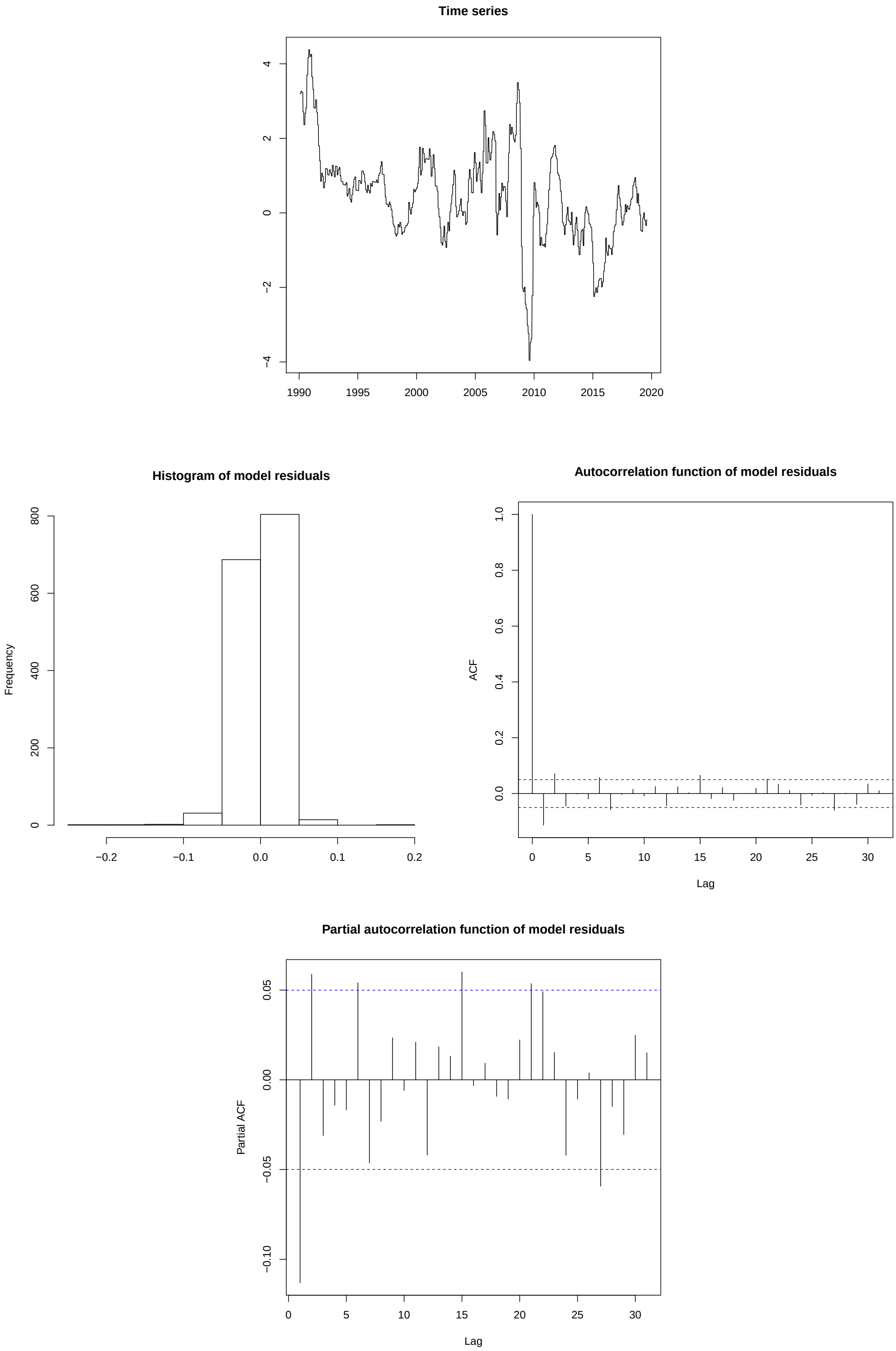

Figure 5: Regression diagnostics: Inflation target: 2%, US

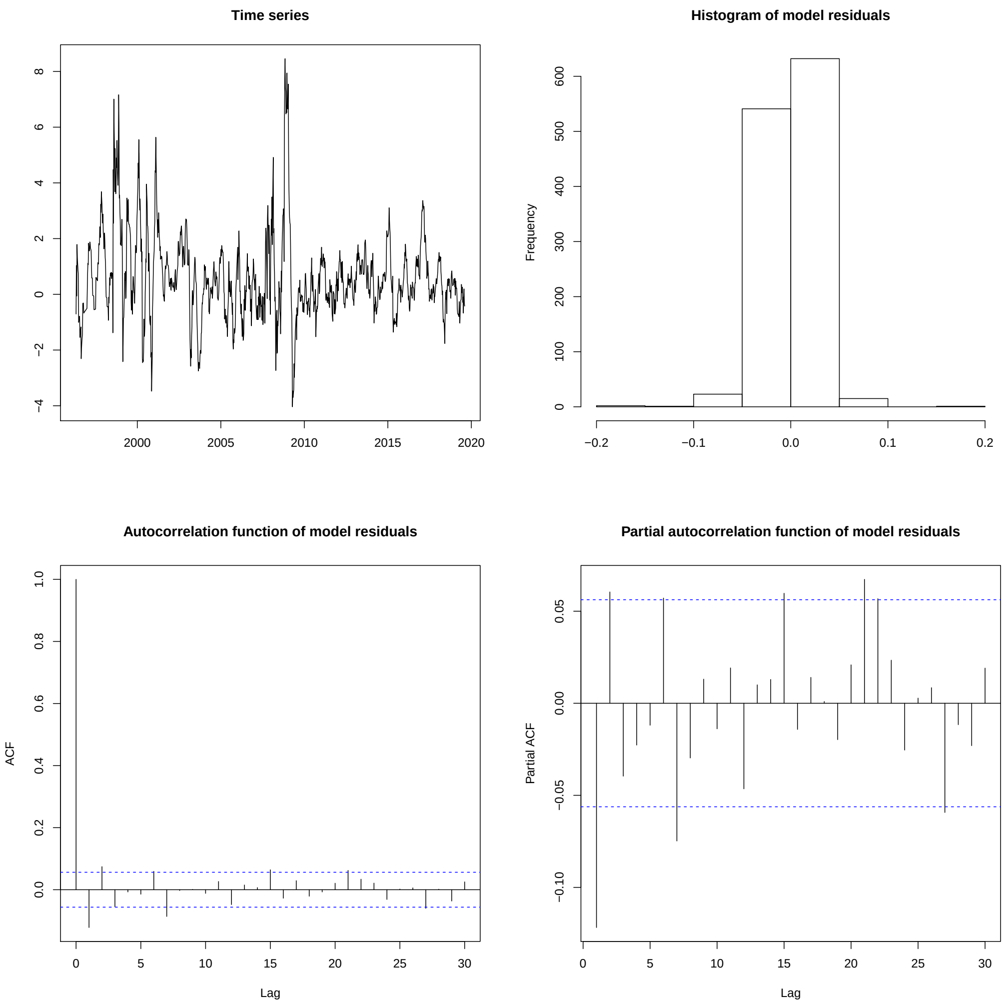

Figure 6: Regression diagnostics: ETF flow, US