Master's thesis

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

Wilhelm Jebsen Mikkelsen
Ruben Ravndal
Martin Tveitstøl

An Early Warning System for Financial Market Corrections on the Oslo Stock Exchange

A Multinomial Logistic Regression Approach

Master's thesis in Industrial Economics and Technology Management
Supervisor: Einar Belsom
June 2019

Foto: Stein Henningsen

NTNU
Norwegian University of Science and Technology

Wilhelm Jebsen Mikkelsen
Ruben Ravndal
Martin Tveitstøl

An Early Warning System for Financial Market Corrections on the Oslo Stock Exchange

A Multinomial Logistic Regression Approach

Master's thesis in Industrial Economics and Technology Management
Supervisor: Einar Belsom
June 2019

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

NTNU
Kunnskap for en bedre verden

Abstract

We use a four-state multinomial logistic regression model in order to estimate the probability of corrections and crises in the Norwegian stock market. The probabilities are subsequently translated into market exposure through a systematic trading algorithm based on the Kelly criterion, aimed at yielding risk-adjusted return in excess of a buy-and-hold strategy in the Oslo Stock Exchange Benchmark Index (OSEBX).

We conclude that financial indicators carry predictive content for the occurrence of market downturns. Particularly, we find that the Price-to-book and Price-earnings multiples, the VIX and the Commodity Channel Index are suitable determinants of stock market development. With a realized Sharpe ratio of 0.86, our candidate strategy outperforms the market at a significance level of 95 %. Thus, we find evidence against semi-strong form of market efficiency in the Norwegian stock market during the period march 2014 - march 2019.

Acknowledgment

We would like to express our sincere gratitude towards our supervisor Einar Belsom. His patient guidance and valuable feedback has been instrumental in ensuring the progress of our work. Einar has provided us with intellectual encouragement and numerous stimulating discussions that have generated the foundation of our report.

Furthermore, we want to express our gratitude towards Chief Strategist Peter Hermanrud at Sparebank 1 Markets. Along with displaying an avid interest in our work, his ideas and insights into financial markets have made an invaluable contribution to our study. A further thanks to Thomas Myrholt for facilitating access to several of the bank's financial variables.

We would also like to thank Sjur Westgaard at the Department of Industrial Economics, and Mette Langaas and John S. Tyssedal at the Department of Mathematical Sciences. Their theoretical understanding of financial time series and logistic regression models have been an indispensable resource in the process of model specification.

Contents

1 Introduction **1**

2 Market Efficiency **4**

2.1 Definition and Degrees of Market Efficiency 4

2.2 Classic Market Anomalies 5

2.3 Market Efficiency on the Oslo Stock Exchange 6

2.4 Sources of Mispricing in Equity Markets 7

3 Economic States **10**

3.1 Crises 11

3.2 Corrections 12

3.3 Pre-correction 14

3.4 Discussion of Correction Definition 16

4 Financial Indicators **17**

4.1 Fundamental Valuation Indicators 18

4.2 Peer Valuation Indicators 19

4.3 Stock Market Characteristics 20

4.4 Omitted Variables 24

4.5 Transformation of Variables 25

5 Model Design **29**

5.1 Existing Early Warning Systems 30

5.2 The Multinomial Logistic Regression Model 31

5.3 Modeling Assumptions 32

5.4 Model Selection: Approach and Criteria 36

6 Systematic Trading Strategies **40**

6.1 Merton's Portfolio Problem 41

6.2 Binary Market Exposure 41

6.3 Gradual Market Exposure: Kelly Criterion 43

6.4 Trading Assumptions 47

7 In-Sample Performance and Model Selection **49**

7.1 In-Sample Performance 50

7.2 Model Selection 52

8 Out-of-Sample Performance **62**

8.1 Trading Results 63

8.2 Assessment of Significance 64

8.3 Discussion 68

9 Conclusion **78**

Appendices **89**

A Crisis and Corrections **90**

B Variables **92**

C Model Selection **104**

D Trading Strategies **106**

List of Figures

3.1 CMAX as crisis indicator on OSEBX 12
3.2 Definition of pre-corrections, corrections and crises 15

4.1 EGARCH trajectories during corrections 21
4.2 Volatility Spike Indicator during the market states 23
4.3 EGARCH during the market states 23
4.4 Price-to-book ratio of OSEBX, with market states 26
4.5 Price-to-book ratio less 1y rolling average, with market states . . 26
4.6 EQNR / 3Y Brent future contract 27

7.1 Probability of tranquil in the training set for Reference model . . 50
7.2 Probabilities in training set for Reference model 51
7.3 Distribution of LPS and Cross Entropy in the validation set . . . 52
7.4 Probability of tranquil in validation set 53
7.5 Probability of tranquil in validation set, with OSEBX 53
7.6 Portfolio values for all models in validation set 55
7.7 Distribution of Sharpe Ratios in the validation set 56
7.8 Marginal effects of indicators on the dependent variable 59
7.9 Portfolio values of all models in validation set 61

8.1 Portfolio values in the test set 63
8.2 Random portfolio values in test set 65
8.3 Distribution of Sharpe ratios for random portfolios 66
8.4 Probabilities of tranquil and pre-correction in test set 68
8.5 Market exposure Candidate strategy 69
8.6 Market exposure relaxed strategy 70

8.7 Return distributions as seen by relaxed strategy 71
8.8 Distribution of Sharpe ratios 73
8.9 Distribution of Sharpe Ratios: CCI 74
8.10 Distribution of Sharpe Ratios: PB250 75
8.11 Distribution of Sharpe Ratios: SX250 76
8.12 PE OSEBX/STOXX less 1y average 77
8.13 Distribution of Sharpe Ratios: VSI 77

B.1 Weighting of EQNR in the OSEBX index 94
B.2 Brent 3Y future and Linked 3Y1M Brent future 95
B.3 Value of OBX as share of OSEBX 97
B.4 Original and transformed Price-to-book ratio of OSEBX 99
B.5 Original and transformed Price-earnings ratio of OSEBX 100
B.6 Original and transformed EQNR / 3y Brent of OSEBX 100
B.7 Original and transformed OSEBX / Synthetic index 101
B.8 Original and transformed PE OSEBX / PE STOXX 101
B.9 Distribution of transformed PB and PE 102
B.10 Distribution of transformed EQ and SX 103
B.11 Distribution of transformed SY and CCI 103
B.12 Distribution of VIX and VSI 103

List of Tables

2.1 Tests of market efficiency of OSE 7

3.1 Average total return during n last days before a correction 14
3.2 Occurrences of the different states 16

4.1 Categorization of correction indicators 25

5.1 Probability notation 32
5.2 Correlation matrix of independent variables 35

7.1 LPS and Cross Entropy of selected models in validation set . . . 52
7.2 Performance of top four models with binary exposure 56
7.3 Performance of top four models with gradual exposure 56
7.4 Coefficients of top performing model in validation set 57
7.5 Performance of Candidate strategy vs. market in validation set . 61
7.6 Coefficients of Candidate strategy 61

8.1 Performance of Candidate strategy with- and ex. market frictions 63

A.1 Occurrences of crises and corrections 91

B.1 Descriptive statistics 102

C.1 Augmented Dickey-Fuller test 105

D.1 Estimated parameters of EGARCH (1,1) volatility model 107

Chapter 1

Introduction

In September 2018, the Price-to-Book (PB) ratio of the Oslo Stock Exchange Benchmark Index (OSEBX) exceeded 2.2 for the first time since the build up to the 2008 global financial crisis. The following months, a price decline of nearly 17 % concluded one of Norway's largest asset pricing bubbles during the last decade. Known as a leading indicator of a financial downturn, could the PB ratio or other financial variables have predicted the pending market correction, and if so, could an informed investor take advantage of this information?

In a perfectly efficient market, systematically outperforming the market is theoretically impossible. Yet, numerous asset pricing anomalies have been discovered over the course of the last decades, challenging the theory of efficient markets[1]. A recent example of market inefficiency is the brief surge in Volkswagen stock price amid the financial crisis of 2008. October 27, Porsche announced its plans to acquire Volkswagen, which led to a short squeeze temporarily making Volkswagen the most valuable company in the world (Allen et al., 2018).

[1] Examples include *the Post-earnings-announcement drift* (Ball and Brown, 1968), *the January effect* (e.g. Rozeff and Kinney Jr (1976), Keim (1983)), *the Day of the week effect* (French, 1980) and *the Size effect* (Banz, 1981)

The occurrences of crises and corrections can have widespread implications for both national and international financial markets and economies. Subsequently, being able to consistently quantify the risk of such occurrences is of great value to investors and policy makers alike. By utilizing the forecasting power of different fundamental financial and monetary variables, Bussiere and Fratzscher (2006), Frankel and Saravelos (2012) and Li et al. (2015) all develop Early Warning Systems (EWSs) for stock market downturns. In contrast to existing literature, which have mainly regarded larger financial crises, we focus on the occurrence of smaller stock market corrections.

The goal of this thesis is to employ a multinomial logistic regression model to assess the forecasting power of several financial indicators on the probability of corrections and crises on the Oslo Stock Exchange. Our hypothesis builds on the presumption that there exist mechanisms inherent in equity markets that may lead to temporary mispricing of assets. We further postulate that there exist financial variables that are able to identify such over- or undervaluation. By categorizing the prevailing market situation into one of four mutually exclusive states, we assume that each state exhibit unique characteristics. If there exist recognizable patterns between the financial indicators and the market states, a suitable classification model should be able to assess the probability of being in each state.

The output of the model is subsequently translated into trading strategies aimed at generating excess return compared to a buy-and-hold strategy in the Oslo Stock Exchange Benchmark Index (OSEBX). We utilize two distinct trading algorithms, the first originating from a binary exposure variable dictated by the expected return of being either fully invested or not at all. The second is a gradual market exposure strategy based on the fractional Kelly criterion. The trading is made as realistic as possible by including market frictions like brokerage commission and bid-ask spread, and by assessing performance on an independent dataset. In turn, the trading results communicate the potential efficiency implications for the Norwegian stock market.

We contribute to existing literature by developing a generic definition for automated classification of stock market corrections. Furthermore, this study presents a novel approach to assessing the forecasting power of financial variables on the probability of market corrections using multinomial logistic regression. By utilizing the output probabilities from the model, we also contribute to the literature on the fractional Kelly criterion by proposing a time-varying, risk-dependent reduction factor. Lastly, our study represent one of few attempts in academic literature to investigate the validity of semi-strong form market efficiency of the Oslo Stock Exchange.

The next chapter presents relevant theory on market efficiency, behavioural finance and several studies concerning the efficiency of OSE. Chapter 3 presents the definitions of corrections and crises, while the various indicators used in our analysis are introduced in Chapter 4. Chapter 5 review existing literature on the classification methods used to predict financial downturns, before presenting the final model specification. The different trading strategies used to convert the probabilities of the multinomial logistic regression model into realized returns are introduced in Chapter 6. Finally, chapters 7 and 8 discuss in-sample and out-of-sample results, along with the potential implications for the efficient market hypothesis at Oslo Stock Exchange.

Chapter 2

Market Efficiency

The logistic regression model introduced in Chapter 5 will culminate into a trading strategy aimed at yielding excess risk-adjusted return compared to a buy-and-hold strategy of the OSEBX index. If successful, the results will have potential implications for the validity of the efficient market hypothesis in the Norwegian stock market. Before proceeding with the modeling specification, the following section presents the different degrees of market efficiency, a selection of classic stock market anomalies, and existing studies concerning the efficiency of Oslo Stock Exchange.

2.1 Definition and Degrees of Market Efficiency

In an efficient market, the price of financial assets reflect relevant information (Dimson and Mussavian, 1998). Following this definition, a direct implication of the efficient market hypothesis is that consistently outperforming the market on a risk-adjusted basis should be impossible.

A stock market can be efficient on either *weak*, *semi-strong* or *strong form.* As the least stringent of the three, the weak form suggests that asset prices reflect all information contained in its historical prices (Malkiel and Fama, 1970). One way of testing this weak form of market efficiency thus involves creating trading rules by utilizing historical prices and assess its ability to consistently beat the market. Examples of such strategies include traditional technical analysis techniques such as using serial correlation of returns to predict future price movements, or buying the previous month's strongest performing stocks under the presumption that the growth will continue.

Market efficiency of semi-strong form states that the price of an asset not only reflect historical prices of the asset itself, but also other publicly available information (Malkiel, 1995). Tests of the semi-strong form may, for example, assess whether prices efficiently adjusts to public information like annual reports or announcements. Lastly, the strongest form of market efficiency implies that asset prices reflect *all* relevant information, including inside information (Malkiel and Fama, 1970). Testing this form can include investigating if the transactions of any informed investors or groups affect price movements.

2.2 Classic Market Anomalies

Throughout the history of capital markets, the efficient market hypothesis has been challenged by various market anomalies. First presented by Ball and Brown (1968), the *Post-earnings-announcement drift* is the tendency of abnormal, positive drift in stock returns in the aftermath of a company reporting surprisingly good results. Rozeff and Kinney Jr (1976) later discovered the *January effect*, a tendency of abnormal high return in January compared to other months. While anomalies often disappear, reverse or attenuate after being documented in academic literature and made publicly known (Schwert, 2003), the *January effect* persisted for many years after its discovery (e.g. Haugen and Jorion (1996); Haug and Hirschey (2006)).

Constituting another calendar rooted anomaly, the *Day-of-the-week effect* is the discovery of significant negative return on Mondays compared to other weekdays (French, 1980). In line with the work of Schwert (2003), however, Philpot and Peterson (2011) conducts a study of more recent research showing that the effect has moved to other days, reversed or vanished.

Lastly, the *Size effect* is the observed trend of smaller firms to exhibit, on average, higher risk adjusted return compared to larger firms. Summarizing prevailing literature on the effect, Van Dijk (2011) finds that the size premium has been positive and large in recent years, and argue that it is premature to conclude that the effect has disappeared.

2.3 Market Efficiency on the Oslo Stock Exchange

The efficiency of Oslo Stock Exchange has primarily been tested on either *weak* or *strong* form, and academic literature on the subject mainly consist of Master's Theses. A summary of these studies is presented in Table 2.1. In general, OSE has proven to be efficient on weak form, but inefficient on strong form.

Using a trading strategy based on support and resistance, Tollefsen (2010) concludes that Oslo Stock Exchange was weakly inefficient during the period 1999 to 2010. Even after controlling for broker commissions and bid-ask spreads, the strategy significantly outperformed a buy-and-hold strategy. Other studies conducting tests of weak form, however, fail to achieve abnormal profit.

Ørpetveit and Hansen (2016) investigated whether OSE was efficient on semi-strong form by using trading strategies based on multiples such as price-to-book (PB) and price-earnings (PE) ratios. Their out-of-sample tests fail to achieve abnormal profit.

Tests of strong efficiency have mainly been conducted by investigating the effect of primary insiders' transactions (e.g. Engevik and Helleren (2009); Zulovic (2012)). The studies all conclude that there are significantly positive returns in the days coinciding with primary insiders purchases, thereby violating the assumptions of a strongly efficient market.

Strategy	Source	Conclusion
Weak form		
Overreaction, buy poor performing stocks	Mamelund (2006)	Y
Earnings announcement premium	Borch (2008)	Y
Support and resistance	Tollefsen (2010)	N
Inter-day trading on momentum	Simonsen (2012)	Y
Trade volume, under- and over-performing	Dalen (2014)	Y
Exponentially moving average	Nyhus and Skjørten (2016)	Y
Pairs trading	Lamache and Sandøy (2016)	Y
Semi-strong form		
Multiples, market capitalization and momentum	Ørpetveit and Hansen (2016)	Y
Strong form		
Transaction of primary insiders' effect on stock return	Engevik and Helleren (2009)	N
Transaction of primary insiders' effect on stock return	Zulovic (2012)	N
Effect of announcement of emissions	Hagestande and Hals (2012)	N
Effect of profit warnings	Sand and Bødal (2013)	N
Transaction of primary insiders' effect on stock return	Hamre and Sande (2014)	N
Effect of announcement of acquisitions	Olafsson and Fossen (2014)	N
Transaction of primary insiders' effect on stock return	Langli (2015)	N
Effect of announcement of acquisitions	Engebretsen (2017)	N

Table 2.1: Summary of market efficiency tests on Oslo Stock Exchange

2.4 Sources of Mispricing in Equity Markets

A general approach towards challenging the validity of the market efficiency hypothesis is to create a systematic trading strategy that consistently outperforms the market on a risk-adjusted basis. In order to do so, a first step is to formulate a hypothesis of the potential anomaly. Our hypothesis builds on the assumption that there may exist temporal and general mispricing in stock markets, and that there are indicators that carry predictive content for such incidents.

Academic literature suggest several different sources of asset mispricing. Many originates in the field of behavioural finance, which seeks to explain the influence of psychology on investors' behaviour and the subsequent effect on asset pricing (Sewell, 2007). During the 1970s and 1980s, several market anomalies were revealed, casting doubt on the perception that asset prices are determined by unbiased expectations of fundamental values. This opened for alternative explanations for why stocks drift away from its fundamental values. Behavioural finance argues that cognitive biases affects the ability to deal with decision-making under uncertainty, and consequently the ability to assess the pricing of stocks.

Drees (2005) highlights the cognitive biases of *overconfidence* and *confirmation* as two well-documented influencers of financial markets. Overconfidence concerns investors' tendency to overestimate own abilities, initially leading to positive short-lag autocorrelation (momentum) and excess volatility. Daniel et al. (1998) claim that "stock prices overreact to private information signals, and underreact to public signals", further arguing the tendency to exaggerate the importance of existing data neglected by others and own ability to generate information (e.g. through interviews, verify rumours and/or analysis). This may cause stock prices to overreact, before being partially corrected if arriving public information contradicts own beliefs. However, if the public information received validate the investor's opinion, further overreaction may be triggered. Continuing overreaction causes momentum in stock prices, which eventually will be reversed by conflicting public news (Daniel et al., 1998). Ultimately, the overconfidence bias leads to negative long-lag autocorrelation (Drees, 2005).

The confirmatory bias is related to belief perseverance; once hypotheses are formed, information that conflicts these hypotheses tend to be ignored, rejected, or misinterpreted (Harmon-Jones and Mills, 1999). This bias is closely linked to overconfidence, since information is processed in a way that reinforces current beliefs, and consequently increases the overconfidence (Rabin, 1998). The confirmatory bias may be particularly prevalent in financial markets, since investors are facing complicated and ambiguous information, and only a selection of information can be processed (Drees, 2005). Drees (2005) incorporates cognitive dissonance and confirmatory bias in an asset pricing model and demonstrates that persistent over- and undervaluation may occur as a consequence.

Another explanation of asset mispricing emphasize the incentive of investors to earn financial return, which according to Brunnermeier and Oehmke (2012) prompts rational investors to buy and hold mispriced stocks simply because they believe that they can sell to a higher price in the future. Barberis et al. (2015) promote this idea, and note that survey evidence suggests that many investors form beliefs of future prices based on extrapolating past price movements. They further develop the Extrapolative Capital Asset Pricing Model (X-CAPM), which investigates stock market dynamics under the assumption that investors are either (1) extrapolative or (2) fully rational. By modeling the theoretical interaction in a stock market consisting of a given fraction of each type of investor, their results show that the model captures "many features of actual prices and returns" (Barberis et al., 2015).

While not a complete list of the factors that influence stock price movements, these cognitive biases all represent examples of why assets prices may drift away from fundamental values. If large or sustained, this discrepancy may in turn trigger financial corrections or crises, forcing the price down to more sustainable levels. The next chapter aims to distinguish periods of overpricing to the periods of ongoing financial turmoil and periods of relative stability. The purpose is to define mutually exclusive states that display distinct characteristics. If appropriately selected, relevant indicators may in turn be able to recognize these characteristics.

Chapter 3

Economic States

We base our analysis on the assumption that the stock market at any given point can be categorized into of four states. The first state represents tranquil periods, where the market is free of financial turmoil. The second state represents a pre-correction period, defined as a given number of days leading up to a market downturn. The implicit assumption is that pre-correction is a period of asset mispricing that subsequently leads to a reversion towards more sustainable price levels. The third and fourth states reflects correction and crisis periods, respectively. While differing in size and duration, both states represent periods of stock market declines.

Statistical classification methods generally aim to identify the values taken by the explanatory variables during different states in order to assign every observation to a state. Thus, the definition of these states determines what the model attempts to recognize, and consequently, may have significant implications on the resulting parameters. The following chapter presents the definitions of crises and corrections and the corresponding distribution of the multinomial dependent variable used to model the probability of stock market downturns.

3.1 Crises

While economists have not yet established consensus on the definition of financial crises (Kaizoji and Sornette, 2008), existing literature present several different approaches. Brunnermeier and Oehmke (2012) refer to bubbles qualitatively as "large, sustained mispricings of financial or real assets". Taking a more quantitative approach, Mishkin and White (2002) consider a percentage decline of at least 20 %. Patel and Sarkar (1998) and Coudert and Gex (2008) consider a CMAX_t ratio defined as the current price level divided by the recent maximum. Both studies use data averaged on monthly frequency and the historical 24-month maximum in order to define financial crisis months.

In line with the work of Patel and Sarkar (1998), we employ the CMAX_t to identify financial crises. However, while the prevailing use involve monthly data, we employ a ratio based on data with daily frequency. Let P_t denote the OSEBX index level at time t, then CMAX_t is defined as:

$$\mathrm{CMAX}_t = \frac{P_t}{\max(P_t, P_{t-1}, ..., P_{t-500})} \tag{3.1}$$

The CMAX_t is subsequently translated into a binary crisis indicator, $I_t^{(crisis)}$, that suggests the presence of a financial crisis if the CMAX_t is more than 1.5 standard deviations below its mean. Equation 3.2 presents the formal definition of the crisis indicator, with the standard deviation denoted by σ. Figure 3.1 illustrates both the CMAX_t and the crisis indicator over the sample period 1983-2019.

$$I_t^{(crisis)} = \begin{cases} 1 & \text{if } \mathrm{CMAX}_t \leq \overline{\mathrm{CMAX}} - 1.5\sigma \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

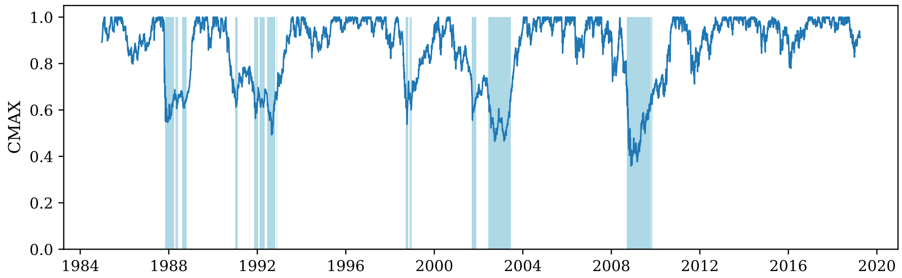


Figure 3.1: CMAX (blue) and crisis indicator (light blue)

3.2 Corrections

To the best of our knowledge, no formal definition of stock market corrections exists in academic literature. According to Schwab Center of Financial Research, the general definition of a market correction is a decline of more than 10 % (Schwab Center for Financial Research, 2018). In a meeting with Chief Strategist Peter Hermanrud at Sparebank 1 Markets, mr. Hermanrud defined corrections as "a decline of at least 8 %, in a relatively short time". He emphasized that in order to be categorized as a correction, the price movement must be large enough, and happen fast enough, to invoke a certain amount of distress in the minds of investors (Hermanrud, 2019).

A preliminary requirement to a working definition of market corrections is that the price level must decline in excess of a given threshold during a given time. In line with the prevailing approach of Sparebank 1 Markets, one of Norway's largest investment banks, we define this threshold to 8 % over 30 trading days.

In order to determine the specific days that constitute a market correction, we proceed by explicitly defining the requirements for any given day to be categorized as the starting point of a correction. In this regard, a requirement is that the OSEBX must be at its maximum level compared to both the previous and following period. This local maximum constraint ensures that the correction begins after a period of growth, so that no correction is defined in the middle of a downturn during a larger crisis. In order to comply with this condition, we stipulate that the price level must exceed that of the previous year, and the following 30 trading days. Thus, the following list summarizes the proposed requirements for a time t to be considered as the starting point of an imminent stock market correction:

- There exists a price decline of at least 8 % within the next 30 trading days
- Today's price level exceeds the level of any of the past 252 trading days
- Today's price level exceeds the level of any of the following 30 trading days

Letting $I_t^{(corr)}$ represent the beginning of a correction at time t, these requirements can be expressed mathematically as

$$I_t^{(corr)} = \begin{cases} 1 \quad \text{if} & \begin{array}{l} \min\{P_{t+1}, ..., P_{t+30}\} < 0.92 \cdot P_t \\ \max\{P_{t-252}, ..., P_{t+30}\} = P_t \end{array} \\ 0 \quad \text{otherwise} \end{cases} \tag{3.3}$$

If a correction start is identified, the correction is defined from top to bottom. With the help of the CMAX-definition, the crises are also defined from top to bottom. When corrections coincide with crises, however, the correction is defined the first six weeks of the crises, and the crises are defined from that point and to the bottom.

3.3 Pre-correction

The rationale behind including a pre-correction period is to identify indicator levels prior to a pending correction. Thus, the length of this period must be determined. In current academic literatre, Bussiere and Fratzscher (2006) and Li et al. (2015) use a period of 12 months in the definition of a pre-crisis period. Given more frequent occurrences and shorter duration, along with the assumption of more temporary mispricing, a shorter pre-period is natural in the case of market corrections.

When determining the length of the pre-correction period, it is important to note that the period leading up to a correction is usually characterized by large growth. Table 3.1 shows the average return during the last 10, 20, 30 and 40 days before the decline. If chosen too long, the pre-correction period may encompass observations which show no real overpricing relative to the levels after the correction. Adversely, a short period may exclude observations that does.

	10 days	**20 days**	**30 days**	**40 days**
Average return [%]	3.9	6.3	9.7	12.8

Table 3.1: Average total return during n last days before a correction

The notion of high returns leading up to the decline is of interest also from a trading perspective. As evident from Table 3.1, an investor that reduces market exposure too early intending to avoid a potential correction might actually bypass a price increase comparable in size to the subsequent decrease. This supports a relatively short pre-correction period.

From a modeling perspective, there is an apparent trade-off between the need of observations for parameter estimation and the desire to only include observations close to the correction start. While a short pre-correction period may increase uncertainty in estimations, it may also help distinguish the indicator levels prior to a correction.

Following these considerations, we define the pre-correction period as the final 20 trading days leading up to a correction. When looking at the ratio between the pre-correction and correction periods, this is comparable to the pre-crisis period proposed by Bussiere and Fratzscher (2006)[1]. Mathematically, the pre-correction indicator $I_t^{(pre)}$ is defined as

$$I_t^{(pre)} = \begin{cases} 1 & \text{if } \exists\ k \in \{1, 2, ..., 20\} \text{ so that } I_{t+k}^{(corr)} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

When combined, the final definitions of crisis-, correction- and pre-correction periods represent the periods imminent or ongoing financial turmoil. The resulting periods bypassed by all definitions thus reflect relative financial stability, and is noted *tranquil*. Figure 3.2 illustrate the different states along with the log price of the OSEBX index. Table 3.2 show the distribution in absolute numbers and in percent of the total sample period.

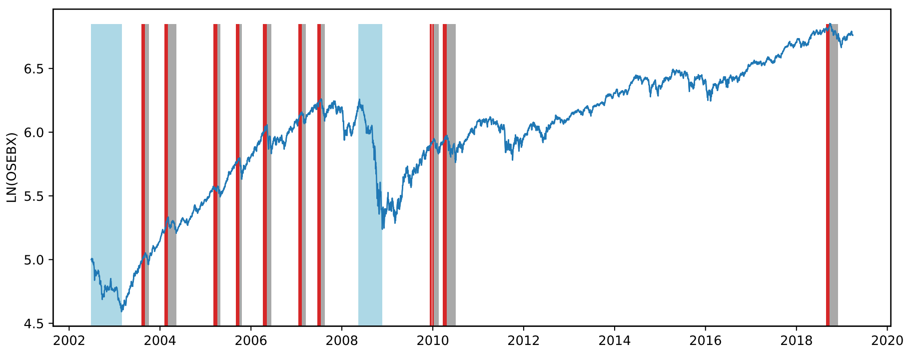


Figure 3.2: Definition of pre-correction (red), corrections (grey) and crises (light blue)

[1]The average crisis lasts 209 trading days, the average correction 25 trading days (Source: Own Calculations)

State	Occurrence [days]	Occurrence [%]
Tranquil	3 408	81
Pre-correction	200	5
Correction	290	7
Crisis	312	7

Table 3.2: Occurrences of the different states

3.4 Discussion of Correction Definition

While the explicit definitions presented in the previous sections are generic and absolute, real world equity markets are far more complex. The boundary between crises and corrections may be unclear, and borderline cases of correction-like behaviour may be wrongfully neglected by a firm definition. Such borderline cases may also have implications for the performance and results of any statistical model based on the classifications.

An apparent type of price movement that is bypassed by our correction definition is the incident of sharp price declines of only near 8 %. When altering the threshold from 8 % to 7 % or 6 %, the number of corrections in the period 2002-2019 increases from 10 to 11 and 14, respectively. A more indirect example may be observed when the OSEBX is drifting moderately downwards before a sharp decline. In this case, the correction conditions may not be satisfied because (1) the decline may have been greater than 8 % but not *within* 30 consecutive trading days, and/or (2) the index was never at a 1-year maximum sufficiently close to the more prominent decline.

While there is no definitive answer on this matter, we note that employing several different definitions may be valuable when used as an input in classification models. As such, comparing the performance under a range of different definitions may enable an assessment of model robustness.

Chapter 4

Financial Indicators

Our objective is to develop a trading strategy based on the estimated probabilities of stock market downturns. Thus, being able to accurately identify variables that are able to convey information on the risk of over- or undervaluation in the market is imperative. Literature focused on creating EWSs for financial crises have utilized a combination of financial ratios and macroeconomic- and monetary variables in order to recognize attributes of economic imbalances. While stock market corrections and more fundamental economic crises differ in both magnitude and frequency, and potentially by cause, we argue that variables such as the Price-to-book ratio may help indicate the overvaluation characterized by all asset pricing bubbles.

This section describes the variables used to construct an Early Warning System for stock market corrections on the Oslo Stock Exchange. We begin by presenting and motivating the three different classes of variables, and discuss their effects from a theoretical perspective. We then proceed with a walk-through of the variables' presumed ability to indicate under- and overvaluation, and describe any transformations applied to them.

4.1 Fundamental Valuation Indicators

The first class of variables is comprised by a set of fundamental-scaled price ratios, which are widely used to relate stock valuation to cash flows, profits or other metrics of underlying value (Lie and Lie, 2002). Examples include the Price-to-book (PB) and Price-earnings (PE) ratios, which reflects the pricing of a company relative to its underlying book value and reported earnings, respectively. The purpose of including these fundamental valuation indicators is to distinguish the sustainable price increases supported by fundamentals from the unsustainable increases spurring asset price bubbles.

According to Herwartz and Kholodilin (2014), one important attribute of such ratios is that trending fundamental drivers of market performance are likely to affect both the numerator and the denominator. If this holds, the ratios themselves are somewhat protected against trends in fundamentals such as inflation and oil price movements. From an econometric perspective, the ratios can thus be viewed as equilibrium relationships cancelling the stochastic drift inherent in stock movements. From a valuation perspective, such stationarity is supported by the notion that the price of an asset is unlikely to exhibit a permanent drift from its underlying value. The idea is that if the numerator and the denominator is affected proportionately, a considerable drift away from the equilibrium is more likely to convey the incident of mispricing.

If a ratio represents an equilibrium relationship, any positive or negative deviation must eventually be accompanied by subsequent mean reversion (Herwartz and Kholodilin, 2014). Either the numerator, the denominator, or both, must adjust in order to restore the equilibrium. Intuitively, market-determined asset prices are more inclined to adjust to underlying fundamentals than vice versa. If this is true, financial ratios may carry predictive or explanatory content for periods of inflated or deflated asset price levels.

We include three fundamental valuation indicators in our analysis. The *Price-to-Book* ratio (PB) is included under the hypothesis that a considerable drift away from underlying equity value suggests increased risk of overvaluation. Similarly, the *Price-Earnings* ratio (PE) conveys overpricing relative to expected future earnings. This is in line with Fu et al. (2019), which conclude that both are suitable indicators of price bubbles. In order to capture the effects of overpricing relative to the oil price, we also include *EQNR/3YBrent* (EQ), the relationship between Equinor and 3-year Brent Oil futures. A more detailed description of the construction of these and the following variables is presented in Appendix B.

4.2 Peer Valuation Indicators

A common approach to stock valuation is to utilize a company's fundamentals directly through calculation methods such as the discounted cash-flow approach in order to obtain the intrinsic value of a stock. Another approach is to use financial multiples and the price of comparable firms in order to indicate a fair price of the company at question. The idea is that if the two companies display otherwise similar characteristics, their price level relative to financial fundamentals should also coincide (Bhojraj and Lee, 2002).

As with the fundamental approach, the ideas of peer valuation can be extrapolated into the analysis of aggregated stock indices. Documented by Markwat et al. (2009), interconnectedness between international capital markets results in volatility contagion between countries. If two indices are interconnected, and their characteristics are resembling, the relationships between them could indicate the incident of over- or undervaluation.

In an attempt to recognize imbalances in the relative pricing of the Norwegian stock market to international markets, our analysis will build on two separate peer valuation indicators. The first is the relationship between the PE ratios of OSEBX and the STOXX Europe 600 index, which reflects the relative pricing of earnings of the companies listed on the two indices. This way, a large increase in the pricing of earnings on the OSEBX relative to the rest of Europe may indicate increased risk of a correction in the Norwegian market. One drawback, however, is the difference in industry composition between the indices, meaning industry-specific events may impact the two indices disproportionately. As a consequence, the indicator may misleadingly signal mispricing when, in reality, the observed change in dynamics is justifiable.

As an alternative, we include the ratio of OBX[1] and a synthetic index developed by Sparebank 1 Markets aimed at replicating the OBX index. When constructing this hypothetical index, each constituent of the OBX index is paired, and subsequently replaced, with a comparable foreign company[2]. Meanwhile, portfolio weights follow the original weights of the OBX, resulting in a synthetic index with the same industry composition as its origin (Hermanrud, 2019).

4.3 Stock Market Characteristics

To complement the relationships of current price levels to underlying fundamentals and the pricing of comparable indices, we consider stock market characteristics with proxies for risk and momentum. These variables are included in order to assess the general investing atmosphere, which may be largely different during the four states.

In stock valuation, expectations of market risk influence the risk premium demanded by investors, and thus affect the appropriate pricing of assets. Intuitively, risk characteristics should also be a suitable measure when identifying the different states of the economy. For example, downward price movements are known to exert stronger impacts on stock volatility in comparison to positive price changes of comparable size (Black, 1976), and studies such as Herwartz and Kholodilin (2014) show that periods of strong price decline go along with considerable increases in stock market volatility.

[1] OBX is assumed to be representative for OSEBX, since it on average has accounted for 88 % of the value of OSEBX. The historical relationship can be observed in Figure B.3.

[2] E.g Equinor and BP, Hydro and Alcoa, Telenor and Telia

We consider three different variables to represent stock market characteristics. The first is the CBOE Volatility Index (VIX), which is the volatility implied by S&P 500 index options. The effect of volatility contagion in global capital markets is documented by Markwat et al. (2009) and Theodossiou and Lee (1993). If OSEBX is indeed affected by movements in foreign capital markets, and the US stock market specifically, the VIX may function as a proxy for expected short-term volatility on the OSEBX.

The second indicator is the Commodity Channel Index (CCI), which is calculated on OSEBX to detect large positive and negative price deviations from its recent mean. The CCI is a popular momentum-based indicator developed by Lambert (1983), and used for predicting stock and index returns by amongst others Kim and Han (2000), Yu et al. (2005), Patel et al. (2015), Kara et al. (2011), and Kordos and Cwiok (2011).

Lastly, the self-developed Volatility Spike Indicator (VSI) is included with the aim of signaling a volatility pattern typically observed during corrections. This is motivated by the notion of correlation between large price declines and increased volatility, as argued by Herwartz and Kholodilin (2014). In order to illustrate the relationship, Figure 4.1 shows the EGARCH volatility trajectories during all corrections in the dataset used to fit the model.

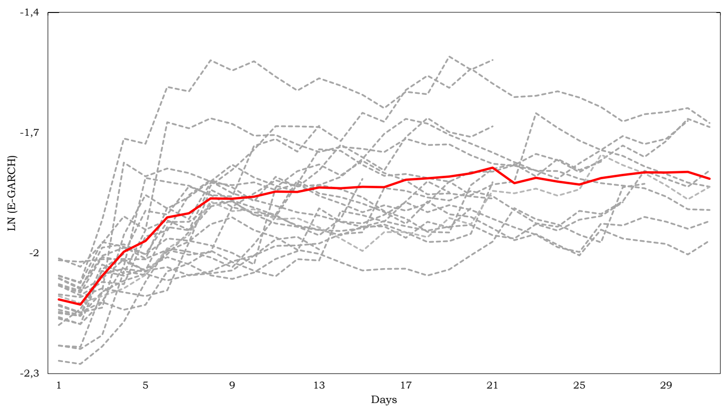


Figure 4.1: EGARCH trajectory during corrections, average value highlighted in red

As evident, a correction start is often characterized by a sharp increase in volatility. Therefore, we design the VSI so as to produce a signal if the EGARCH volatility increases in excess of a certain threshold. Similarly, a correction end signal is produced if the ratio is *less* than some threshold, meaning the volatility has started to decrease significantly. The signal is dampened until EGARCH has reached below some threshold value. The goal is to design an indicator that signals the sharp volatility increases characterizing correction periods, while minimizing signals produced outside corrections and crises. The VSI is defined as follows, where EG_t is the estimated EGARCH volatility at time t:

$$\mathrm{VIS}_t = \mathrm{EG}_t \cdot \mathrm{UP}_t \cdot \mathrm{DOWN}_t \tag{4.1}$$

$$\mathrm{UP}_t = \begin{cases} 10 & \text{if } \frac{\mathrm{EG}_t}{\mathrm{EG}_{t-2}} \geq 1.4 \\ 10 & \text{if } \mathrm{UP}_{t-1} = 10 \ \wedge \ \mathrm{EG}_t \ \geq 0.01 \\ 1 & \text{otherwise} \end{cases} \tag{4.2}$$

$$\mathrm{DOWN}_t = \begin{cases} \frac{1}{10} & \text{if } \frac{\mathrm{EG}_t}{\mathrm{EG}_{t-2}} \leq 0.85 \\ \frac{1}{10} & \text{if } \mathrm{DOWN}_{t-1} = \frac{1}{10} \ \wedge \ \mathrm{EG}_t \ \geq 0.009 \\ 1 & \text{otherwise} \end{cases} \tag{4.3}$$

The result of this definition is shown in Figure 4.2, where VIS is plotted against tranquil, pre-correction (red), correction (grey) and crises (light blue) in the period used for training of the model.

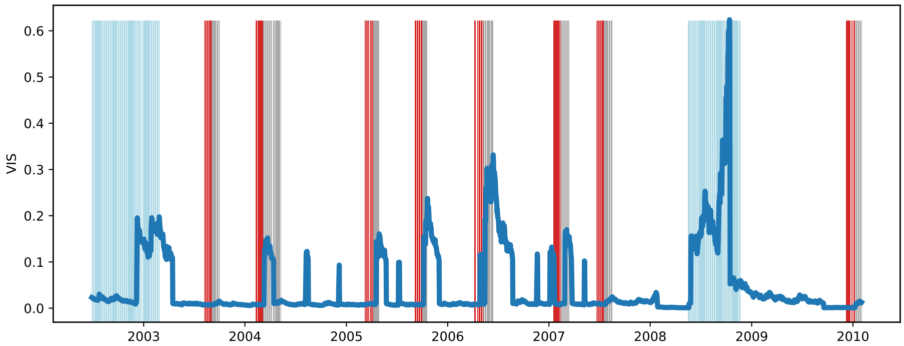


Figure 4.2: Volatility Spike Indicator during the market states

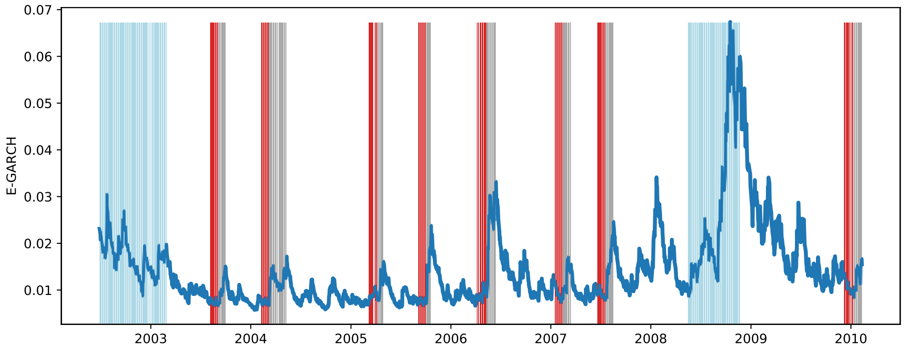


Figure 4.3: EGARCH during the market states

By comparing VIS (figure 4.2) and the EGARCH (figure 4.3), the indicator levels are more distinct in the former case. This may help the model in categorization, given that the spikes correspond with the corrections. However, there are corrections where VSI does not produce signals, such as in late 2003 and in the middle of 2007. There are also cases where VSI spikes without a correction taking place, such as in 2004. These cases may confuse the model in terms of what characterizes the different states.

4.4 Omitted Variables

The aforementioned indicators hardly constitute an exhaustive list of variables that affect price movements in the Norwegian stock market. Amongst the most prominent omitted indicators are monetary- and macroeconomic variables such as term spread, unemployment rate and real per capita GDP. Previous studies such as Li et al. (2015) and Herwartz and Kholodilin (2014) have found that these are all sound indicators of asset price bubbles.

An important distinction is that while existing literature have centered around predicting stock market crises, we focus our analysis towards the more frequent and less dramatic short term corrections. Monetary- and macroeconomic variables such as interest rate or the unemployment rate change slowly, and are associated with larger macroeconomic trends and potential economic imbalances. Moreover, low sampling frequency imply long chains of static values when considering daily observations. Due to this lack of responsiveness, we expect that such variables will be less effective in identifying more temporary asset mispricing. This is further supported by the findings of Herwartz and Kholodilin (2014) which, in a study predicting stock market bubbles, conclude that "financial ratios are uniformly most relevant for modeling and prediction of periods of excess stock market valuation" (Herwartz and Kholodilin, 2014).

Other common financial valuation multiples include EV/EBITDA, EV/Sales and Earnings/Dividend. The exclusion of these variables are primarily the result of data availability limitations. The variables used in further analysis thus consist of a selection of fundamental- and peer valuation indicators along with indicators of market risk. Table 4.1 summarizes the final variable selection.

Indicator	Measure
Fundamental valuation	
PB	Price-to-book ratio of OSEBX
PE	Price-earnings ratio of OSEBX
EQNR/3Y BRENT (EQ)	Ratio of EQNR and 3Y Brent contract
Peer valuation	
OBX/Synthetic (SY)	Industry-neutral pricing of OBX
PE OSEBX vs. STOXX (SX)	Pricing of OSEBX relative to Europe
Stock market characteristics	
VIX	Implied volatility of S&P 500
CCI	Momentum metric
Volatility Spike Indicator (VSI)	Volatility spikes on OSEBX

Table 4.1: Categorization of correction indicators

Omitted Variable Bias

Omitting the aforementioned variables could impact our model in other ways than failing to capture their marginal effects on the probabilities of the different states. If an important factor which is correlated with the other explanatory variables is omitted, its effects will be somewhat attributed to the included covariates, causing the resulting estimators to be biased (Wooldridge, 2015). This substantiates the importance of assessing model performance on an independent dataset.

4.5 Transformation of Variables

Compared to larger financial crises, we assume that corrections are caused by smaller, temporary mispricings. As discussed in Section 2.4, overvaluation may be caused by cognitive biases leading to momentum, which eventually will be reversed. Because of the short duration and high frequency of corrections, the variables presented in the previous sections are not necessarily suitable indicators of corrections in their original form.

One example is the Price-to-book ratio, whose development is illustrated in Figure 4.4. In 2008, the PB ratio reached unprecedented levels in the preface to the global financial crisis, and has later been recognized as an important indicator of financial crises in general (e.g. Herwartz and Kholodilin (2014), Fu et al. (2019)). The OSEBX experienced a new price correction in 2010, but since the PB levels had not yet recovered from the crisis, the relatively modest PB ratio would hardly signal an impending downturn.

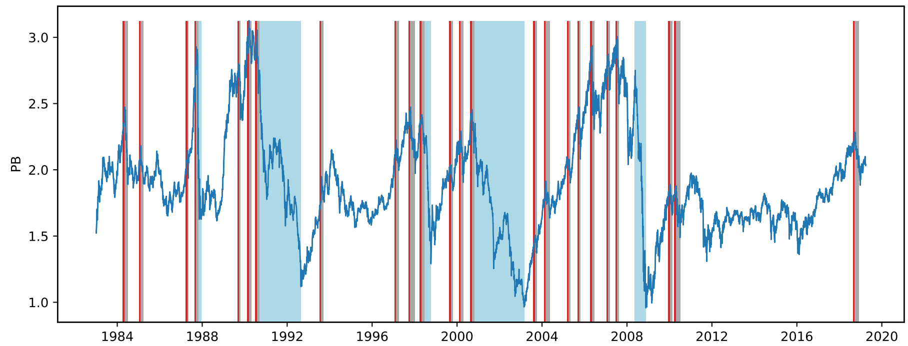


Figure 4.4: Price-to-book ratio of OSEBX from 1983 - 2019, with market states

A rolling average may be suitable for identifying cases where the market pricing has increased too fast relative to fundamental drivers and peers. Figure 4.5 shows the PB ratio less its 1-year rolling average.

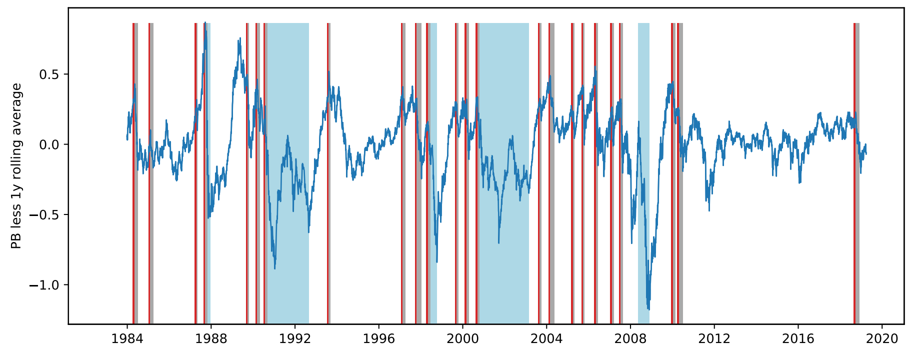


Figure 4.5: Price-to-book ratio less 1y rolling average, with market states

Comparing Figures 4.4 and 4.5, the latter is clearly more suited for indicating an overly rapid increase. This supports the idea that, for some variables, temporary asset mispricing is better recognized by the short-run optimism reflected in the discrepancy from a recent mean.

Another important notion is that we aim to investigate the typical levels of each independent variable during any given state of the dependent variable. If an independent variable drifts over time, however, its levels will be largely time-dependent, making the relationship between the states and indicator level unclear. One example is the ratio between ENQR stock price and the Brent futures contracts. Following the significant drop in oil price in 2014, EQNR initiated substantial cost saving programs which in 2017 was reported to have resulted in USD 3.2 billion annual efficiency gains (Hovland, 2017), thereby contributing to the justification of a higher price multiple. As displayed in Figure 4.6, this relationship shows a clear drift.

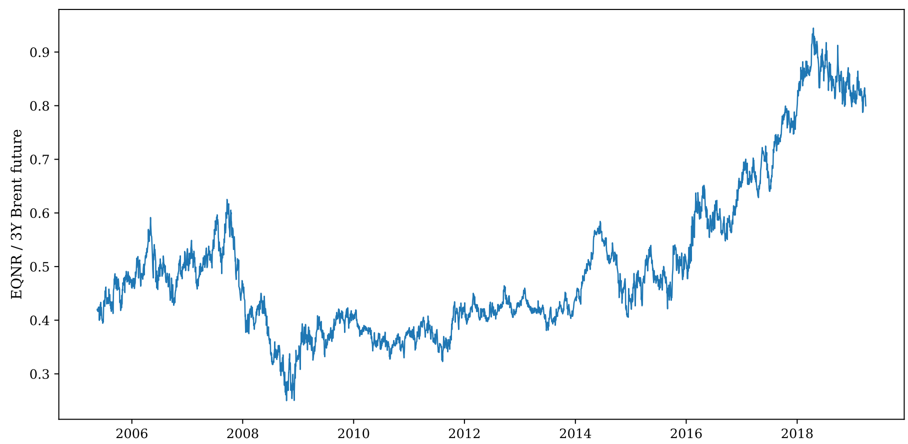


Figure 4.6: EQNR / 3Y Brent future contract

Following these considerations, all fundamental and peer valuation indicators are transformed by subtracting a rolling average. Letting X_t represent the original variable, we obtain the transformed version x_t as indicated by Equation 4.4. The rationale is to create variables that are better suited to represent the short-term optimism characterized by market corrections, in addition to eliminate any non-stationarity in the original variables. The variables related to stock market characteristics are used in their original form.

$$x_t = X_t - \frac{1}{\tau}\sum_{j=1}^{\tau} \mathrm{X}_{t-j} \quad , \quad \forall\ X_t \in \{\mathrm{PB}_t, \mathrm{PE}_t, \mathrm{EQ}_t, \mathrm{SY}_t, \mathrm{SX}_t\} \tag{4.4}$$

When determining the time window for the moving average, the interval should be long enough to enable distinct deviations from the mean, but short enough to exclude effects of the previous correction or crisis before entering a new one. With regards to the number of different lags to consider, a trade-off must be made between testing a representative sample and the increasing risk of randomness; if numerous lags are included, some of them may by chance happen to correspond well with the corrections. Following these considerations, we include the indicators less both their 120 and 250 trading days rolling average in further analysis. All indicators are plotted in Appendix B.

Chapter 5

Model Design

Chapter 4 presented the different indicators that are expected to impact the probabilities of being in each of the states introduced in Chapter 3. In order to quantify the impact of the indicators, we now turn to modeling the probabilities of being in each of these states. This section first describe existing literature on early warning systems, before presenting and justifying the final choice of classification model along with necessary assumptions.

The classification model chosen is later used to test various indicator compositions and trading rules. This chapter will cover the methodology of the model selection process and the specific metrics used for choosing a robust variable composition and trading strategy.

5.1 Existing Early Warning Systems

Since the 1990s, several empirical studies have constructed early warning systems for financial crises, mainly adopting one of two approaches. Kaminsky et al. (1998) introduced a static signal extraction approach, which involve monitoring variables that are expected to display atypical behaviour in the period leading up to a financial crisis. The model is designed to signal increased risk of an impending crisis if these variables exceed a threshold value, determined as a given percentile of each indicator's sample distribution. Continuing the concept of signal extraction, Casu et al. (2011) later developed a dynamic approach to defining the threshold value. This model suggests increased risk if any indicator exceeds a rolling average by a pre-defined number of standard deviations.

As an alternative to the signal extraction approach, Frankel and Rose (1996) proposed the use of binary logit and probit regression models to examine currency crises. Manasse et al. (2003) and Fuertes and Kalotychou (2006) similarly utilized logistic regression in the context of debt crises in emerging markets. Manasse et al. (2003) further argued that logit should be used instead of probit models when the dependent variable is unevenly distributed amongst its possible values. As periods of crises and corrections occur rather seldom compared to periods without them, this is clearly the case when predicting financial downturns.

Economic turmoil during crises may force the indicators into an adjustment process before exhibiting more sustainable behaviour. As argued by Bussiere and Fratzscher (2006), the signaling indicators can thus be reasonably expected to take differing values during crises and otherwise tranquil periods. In order to avoid this potential issue when using a binary dependent variable, several authors either introduce a dummy variable to allow for alternative coefficients in crisis periods (e.g. Peter (2002), Manasse et al. (2003)) or eliminate the crisis observations from the sample entirely (e.g. Fuertes and Kalotychou (2006), Savona and Vezzoli (2015)). Bussiere and Fratzscher (2006), however, propose the use of a three-state multinomial dependent variable. The results showed that utilizing a multinomial regression model constituted a 'substantial improvement' in predicting financial crises compared to an otherwise equivalent binomial model.

Less prevalent methods include k-means clustering, a data-driven approach which involves partitioning the data into different clusters based on maximizing within-cluster similarity and between-cluster disparity. Fuertes and Kalotychou (2006) employed this approach, although concluding that the k-means approach was outperformed by a binary logit specification. Fioramanti (2008) proposed the use of artificial neural network methods which, under certain conditions, outperformed more traditional methods. An apparent drawback, however, is the non-parametric nature of the approach, meaning that interpreting the marginal effects of each variable is challenging. Consequently, Fioramanti (2008) argue that the neural network approach delivers limited value to policy makers.

We consider a logistic regression model in order to model the probability of financial downturns in the Norwegian stock market. Compared to neural network approaches, a useful attribute of the logit specification is its parametric character, enabling interpretation of the marginal effect of the different covariates. As Bussiere and Fratzscher (2006), we consider a multinomial specification, but expand to a four-state dependent variable which include the occurrence of smaller market corrections.

5.2 The Multinomial Logistic Regression Model

We employ a four-state multinomial logistic regression model based on the four states identified in Chapter 3 . The first state (Y_t=0) represents tranquil periods, where the market is free of financial turmoil. The second state (Y_t=1) represents a pre-correction period, defined as 20 trading days before the decline. The third and fourth states reflect correction (Y_t=2) and crisis (Y_t=3) periods, respectively.

The idea behind the multinomial logistic regression model is to estimate a set of weights that, when linearly combined with the explanatory variables, can be transformed into a set of probabilities of being in each state at a given time. Letting $\pi_{it} := \mathrm{P}(Y_t{=}i)$, Equation 5.1 display the multinomial logistic link function, while Equation 5.2 presents the resulting expressions when solving for the desired probabilities. The parameters are estimated through Maximum log likelihood, and the logistic regression is executed using the Sklearn package in Python.

$$\eta_{it} = \ln\left(\frac{\pi_{it}}{\pi_{0t}}\right) = \alpha_i + \boldsymbol{\beta}_i \boldsymbol{x}_t \ , \quad \forall\ i \in \{1, 2, 3\} \tag{5.1}$$

$$\pi_{it} = \frac{\exp(\eta_{it})}{\sum_{j=0}^{3} \exp(\eta_{jt})} \ , \quad \forall\ i \in \{0, 1, 2, 3\} \tag{5.2}$$

Notation	Probability of being in
π_{0t}	Tranquil
π_{1t}	Pre-correction
π_{2t}	Correction
π_{3t}	Crisis

Table 5.1: Probability notation

5.3 Modeling Assumptions

Logistic regression distinguishes itself from linear regression in the sense that it relaxes several of the key assumptions made by ordinary least square based models. Logistic regression does not require a linear relationship between dependent and explanatory variables, nor is it required that the residuals are normally distributed. Furthermore, homoscedasticity is not required. There are, however, some assumptions that still apply.

Independent Observations

Logistic regression assumes that observations are independent, meaning that one observation does not affect another. Violation of independence have implications for the derivation of the log likelihood function, as the assumption $P(Y_t{=}i|Y_{t-1}, ..., Y_1) = P(Y_t{=}i)$ does not hold[1]. In that case, the model coefficients will no longer be efficient, and statistical inference will in turn lead to incorrect conclusions since the standard deviations of the coefficients are underestimated (Schreiber-Gregory, 2018).

Intuitively, the probability of experiencing a crisis or correction any given day will be highly dependent on the situation the previous day. The dependent variable described in Chapter 3 can thus be reasonably expected to exhibit serial correlation. If the dependent variable is indeed serially correlated, there is explanatory power in the previous outcomes. If untreated, this explanatory power will end up in the error term and subsequently cause serially correlated error terms. Consequently, the standard deviations of the independent variables' coefficients will be artificially small.

A possible measure to decrease the serial correlation is to include previous realizations of the dependent variable as explanatory variables. In similar studies of logistic regression models, this approach is employed in both Flahaut (2004), which models road safety, and Atkinson and Massari (2011), which model land sliding. An important notion, however, is that the dependent variable described in Chapter 3 is inherently distinct from those of both Flahaut (2004) and Atkinson and Massari (2011). To illustrate, consider the choice of an investor the first day into a market correction. While the next day will also be classified as being in a correction period, this information is not known to the investor. The dependent variable is determined *ex-post*, and therefore, cannot be used in prediction.

[1]See Appendix C for the derivation of the MLL estimator

Absence of Multicollinearity

Another assumption is that the explanatory variables exhibit little or no multicollinearity (Schreiber-Gregory, 2018). If violated, multicollinearity may introduce unreliable coefficients (Kleinbaum et al., 2002). Table 5.2 display the correlation between the variables used in the logistic regression model, which show moderate to high correlation amongst several indicators. As many of the variables are mathematically interconnected, this is hardly surprising. For example, both PB, PE and SY have the same numerator (OSEBX price level), meaning an isolated price movement will affect all indicators proportionately.

According to Gujarati and Porter (2003), multicollinearity does not affect the performance when extrapolating the model to new data, provided that the multicollinearity patterns are the same for both datasets. Since the final assessment of our model is performed on an independent dataset, the effects of multicollinearity are partially mitigated assuming that the pattern remains constant throughout our entire data sample. A large sample size will also reduce the problem by producing more precise parameter estimates (Kleinbaum et al., 2002).

Large Sample Size

Due to the use of Maximum log likelihood estimation, logistic regression usually requires a relatively large sample size. According to Schreiber-Gregory (2018), a common guideline is that the occurrences of the least frequent state should exceed ten times the number of independent variables. With eight independent variables, the state with the least number of observations need at least 80 observations. There are 160 observations of pre-correction in the data sample used to fit the model. Therefore, the sample size should not pose a major issue.

	PB250	PB120	PE250	PE120	EQ250	EQ120	SX250	SX120	SY250	SY120	VIX	CCI	VSI
PB250	1.00												
PB120	0.88	1.00											
PE250	0.71	0.82	1.00										
PE120	0.47	0.76	0.84	1.00									
EQ250	0.69	0.68	0.77	0.65	1.00								
EQ120	0.48	0.60	0.60	0.68	0.88	1.00							
SX250	0.60	0.47	0.65	0.45	0.56	0.34	1.00						
SX120	0.59	0.59	0.69	0.64	0.63	0.54	0.89	1.00					
SY250	0.82	0.68	0.42	0.18	0.43	0.24	0.36	0.28	1.00				
SY120	0.70	0.78	0.59	0.48	0.47	0.39	0.34	0.35	0.85	1.00			
VIX	-0.83	-0.71	-0.46	-0.30	-0.42	-0.26	-0.59	-0.52	-0.75	-0.58	1.00		
CCI	0.78	0.87	0.75	0.74	0.64	0.61	0.48	0.61	0.52	0.56	-0.63	1.00	
VSI	-0.32	-0.39	-0.50	-0.46	-0.39	-0.30	-0.33	-0.27	-0.11	-0.21	0.18	-0.42	1.00

Table 5.2: Correlation matrix of independent variables

5.4 Model Selection: Approach and Criteria

We seek to identify the model with the lowest expected prediction error over an independent dataset. This is because the model performance must be evaluated on a new dataset for enabling an assessment of its ability to generalize. With eight indicators (some of which include separate lags) and different trading rules, there is a wide selection of 1943 potential variable compositions.

The majority of existing literature on EWSs for financial crises emphasize in-sample testing. However, in a study of prevailing models, Berg et al. (2005) argues that the EWSs performed largely unsatisfactory in out-of-sample tests. In a study spanning a cross section of both developed and emerging markets, Herwartz and Morales-Arias (2009) also concludes that in-sample return predictability does not necessarily imply ability to predict out-of-sample returns.

An apparent drawback with in-sample testing is that the prediction ability strictly increases with model complexity (Friedman et al., 2001). Consequently, adding one variable will always reduce the training error, and model selection will be biased towards specifications encompassing many variables. This increased complexity may in turn cause overfitting, leading to high estimation error and models that may generalize poorly on independent datasets.

To better identify the model with the best ability to generalize onto an out-of-sample set, we follow the approach of Friedman et al. (2001) and split the original dataset into (1) a training set, (2) a validation set and (3) a test set. In the training set, the different models are fitted and the parameter estimates are generated. The validation set is subsequently used to estimate the out-of-sample performance across all models using a set of predefined metrics. Based on the performance in the validation set, a final variable composition and trading strategy is selected for assessment in the test set.

A common approach is to divide the dataset so that the training, validation and test sets constitute approximately 50 %, 25 % and 25 %, respectively (Friedman et al., 2001). Data availability is a constraining factor, and there is a balance between having sufficient data for training, validation and testing. An increased length of the training set will make the parameter estimates more reliable, while increasing the validation period will reduce the randomness and put more confidence in the model selection. Increasing the test period will similarly decrease the randomness of the ultimate results by better assessing the generalization performance of the model (Friedman et al., 2001).

Given that the sample size is well within the requirements posed by the log likelihood estimation, we reduce the training set at the benefit of a longer test period. The complete dataset, ranging from 28.06.2002 to 29.03.2019, is first divided into 30 % used for testing, and 70 % for training and validation. The 70 % are thereafter divided into 65 % used for training, and 35 % used for validation. We then end up with a training set from 28.06.2002 to 08.02.2010 (45.5 %), validation set from 09.02.2010 to 17.03.2014 (24.5 %) and test set from 18.03.2014 to 29.03.2019 (30 %).

To avoid overfitting, no model is allowed to encompass both lags of the same variable. However, the model is allowed to incorporate different lags for different indicators.

Performance Metrics

The performance metrics are used for model selection in the validation period and to assess the ultimate performance in the test set. Thus, selecting metrics that are able to appropriately identify models with desirable generalization ability is essential. In doing so, we propose a combination of statistical and financial metrics. The statistical metrics measure to what extent the probabilities estimated conform with the actual outcomes, while the financial metric measures the risk-adjusted return.

Two statistical metrics are chosen. The first, known as Logistic Probability Loss, is commonly used to evaluate the overall performance of different classifiers (Friedman et al., 2001). This metric essentially measures the average deviation between the probability and their expected values (0 or 1), and is expressed in Equation 5.3. y_t and p_t denotes the actual value of the binary dependent variable corresponding to a given class and the probability resulting from the model, respectively.

$$LPS = -\frac{1}{T}\sum_{t=1}^{T} y_t \ln(p_t) \tag{5.3}$$

An apparent drawback of the LPS metrics is that it only measures the performance of a single class. Although we are principally concerned with the models ability to predict the tranquil state, cross entropy is used to supplement the single class evaluation. This is motivated by the ability to quantify the overall performance as well as the fact that cross entropy is a common method when comparing classification algorithms (Friedman et al., 2001).

$$\text{Cross Entropy} = -\frac{1}{T}\sum_{i=0}^{3}\sum_{t=1}^{T} y_{it} \ln(p_{it}) \tag{5.4}$$

Neither LPS nor Cross Entropy weigh the consequences of committing the estimation errors. Since the absolute value of the average return during corrections and crises are higher than in the tranquil periods, failure to predict turmoil is more severe than failure to predict tranquility. Furthermore, the reliability of the metrics are largely conditional on the accuracy of the dependent variable definition. The models are punished for reducing the probability of tranquility during price declines which are not defined as a correction or crisis - even though it may be correct. Considering the discussion on the sensitivity on the definition in of the dependent variable in Section 3.4, this is a serious concern.

In order to choose a model and assess its performance relative to the market, we employ the Sharpe Ratio to proxy risk-adjusted return. First developed by William F. Sharpe in 1966, the ratio is a measure of an investment's return in excess of the risk free-return per unit of risk (Sharpe, 1966), and is widely considered as an important tool for comparing portfolios (Ledoit and Wolf, 2008). An important attribute of the Sharpe ratio is that it punishes estimation errors according to the real consequences represented by risk and return. The explicit formula is presented in Equation 5.5, where $\overline{r}_i$ and σ_i denote the realized annual return and volatility.

$$S = \frac{\overline{r} - r_f}{\sigma} \tag{5.5}$$

Chapter 6

Systematic Trading Strategies

The output of the multinomial regression model presented in Chapter 5 is a set of probabilities that reflect the risk of stock market downturns. The aim of the trading algorithms presented in the following chapter is to translate these probabilities into a dynamic market exposure that, in turn, will yield excess risk-adjusted returns. In essence, we consider an optimal portfolio allocation problem with one risky and one risk-free asset. At any given point, the investor must decide what fraction of wealth to invest in the stock market, and what fraction to leave out.

As a preparatory measure, the chapter introduces Merton's portfolio problem and the role of utility functions when solving the problem. As a preliminary trading suggestion, we introduce a binary exposure strategy based on the expected return of the stock market. Subsequently, we present a trading strategy based on gradual market exposure and the fractional Kelly criterion. Lastly, the chapter summarizes the trading assumptions governing the realization of returns under the different trading algorithms.

6.1 Merton's Portfolio Problem

Formulated and solved by Robert C. Merton in 1969, Merton's portfolio problem is a well-known problem in continuous-time finance and inter-temporal portfolio choice. The problem considers an investor that must decide how much to consume and how much to invest, the latter distributed amongst a risky and a risk-free asset. The solution to the problem stems from the desire of maximizing the investor's utility, and the procedure will thus depend on the investor's risk preference (Merton, 1969).

The concept of utility was first introduced by the dutch mathematician Daniel Bernoulli as early as 1738, who noticed that under uncertainty, people did not always act as to maximize monetary gain (Bernoulli, 1738). Bernoulli noted that while there existed a direct link between expected wealth and utility, the marginal increase in utility diminished as the monetary gains increased. The findings laid the foundation of the economic theory of risk aversion, risk premiums and utility.

6.2 Binary Market Exposure

A novel trading strategy suggestion can be obtained by considering the preferences of a risk-neutral investor. Faced by the choice of allocating wealth between a selection of assets, such an investor would invest exclusively in the asset with the highest expected return, completely ignoring its risk profile relative to that of the other assets (Markovitz, 1959). This is analogous to the preferences of an investor with a linear utility function $u(\mathbf{w}) = \mathbf{w}\bar{\mathbf{r}}$, where $\mathbf{w}$ denotes the weightings in each asset and $\bar{\mathbf{r}}$ their expected returns. Consequently, optimizing utility with respect to $\mathbf{w}$ would mean full exposure to the asset with the largest expected return.

Analogous to Merton's portfolio problem with no consumption, we now consider the problem of deciding how to allocate wealth between the OSEBX and a risk-free asset yielding constant rate of return r_f. Making use of the output from the logistic regression model, we compare the expected return of full and no market exposure by utilizing the estimated probability $\hat{\pi}_{0t}$ of being in a tranquil state. Let $\overline{r}_0$ denote the average return of the tranquil periods, and $\overline{r}_{2,3}$ the average return during corrections and crises. Based on the alternative that yields the highest expected return, the risk neutral investor will determine the binary exposure variable $f_t^{(bin)}$ accordingly:

$$f_t^{(bin)} = \begin{cases} 1 & \text{if } \hat{\pi}_{0t} \cdot \overline{r}_0 - (1 - \pi_{0t}) \cdot \overline{r}_{2,3} > r_f \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

Effectively, this trading strategy creates a static probability threshold that, following Equation 6.1, will dictate full market exposure if $\hat{\pi}_{0t} > \frac{r_f + \overline{r}_{2,3}}{\overline{r}_0 + \overline{r}_{2,3}}$. Using the average returns in the training period, the threshold amounts to 0.69.

Naturally, this threshold is critical for the eventual performance of such trading strategies. Given the nature of the binary exposure, a small change in the probability estimate may cause a large impact on the resulting market exposure. Thus, the strategy heavily relies on the estimates of $\overline{r}_0$ and $\overline{r}_{2,3}$. Alternative trading strategies may build on gradual market exposure, which might decrease the information loss incurred when transforming the input probabilities into a binary variable. Moreover, it may alleviate the potential problem of high threshold sensitivity, increasing the robustness of the model. On the other hand, frequent rebalancing will incur higher level of transaction costs which might offset the potential gain.

6.3 Gradual Market Exposure: Kelly Criterion

Another trading strategy evolves from the desire to maximize logarithmic utility. Originating in the world of gambling, the Kelly criterion is well-known as a formula to calculate the optimal bet size. According to the criterion, one should always bet the fraction of wealth that maximizes the expected logarithmic growth rate (Peterson, 2017).

This idea can similarly be applied to the portfolio allocation problem faced in capital markets. Consider the portfolio P consisting of a fraction f in a risky asset S and fraction $(1-f)$ in a risk-free bond B. The risky asset is assumed to follow Geometric Brownian Motion (GBM), with an expected drift μ and volatility σ, while the bond pays a constant rate r. The processes followed by the risky asset, the bond and the portfolio, is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{6.2}$$

$$dB_t = rBdt \tag{6.3}$$

$$dP_t = (f\mu + (1-f)r)\, P_t dt + (f\sigma)P_t dW_t \tag{6.4}$$

where W_t denotes a Wiener process. Assuming that P_t is log-normally distributed with drift $\mu_P = f\mu + (1-f)r$ and volatility $\sigma_P = f\sigma$, and we write

$$\frac{P_t}{P_0} = \exp\left[(f\mu + (1-f)r)\, t + (f\sigma)\sqrt{t}Z_t\right] \tag{6.5}$$

where $Z_t \sim \mathcal{N}(0,1)$. Now, the Kelly criterion states that maximizing investor utility is equal to maximizing the expected logarithmic growth rate. By applying $\mathbb{E}\left[\ln(\cdot)\right]$ to both sides of Equation 6.5, we obtain the expected logarithmic growth rate, $G(f)$, of the portfolio

$$G(f) = \mathbb{E}\left[\ln\left(\frac{P_t}{P_0}\right)\right] = \left(f\mu + (1-f)r - \frac{(f\sigma)^2}{2}\right)t \tag{6.6}$$

We can now maximize the expression with respect to market exposure, f, by solving $\frac{\partial G}{\partial f} = 0$, which (for $t > 0$) yields:

$$f^* = \frac{\mu - r}{\sigma^2} \tag{6.7}$$

The benefits of the Kelly criterion has been widely documented in academic literature (eg. Thorp (2011), Ziemba et al. (2003) and Peterson (2017)). Applied in the field of portfolio optimization, the resulting money management strategy has proved superior to comparative strategies in many aspects. In particular, it maximizes expected growth rate and the median of terminal wealth, the latter of which is especially useful in capital markets where asset returns often exhibit highly skewed distributions (Nekrasov, 2014).

Theoretically, using the exposure indicated by the Kelly criterion will yield the highest expected log return (Ziemba et al., 2003). However, an important notion is that while μ and σ must be approximated, the optimal fraction presented in Equation 6.7 ignores the uncertainty in these estimates. In the presence of such uncertainty, Baker and McHale (2013) showed that reducing the investment in the risky asset improves the expected utility. The resulting *fractional* Kelly strategy is less profitable, but also carry less risk (Nekrasov, 2014).

A common approach in existing literature is to simply adjust the Kelly fraction using a fixed constant (Baker and McHale, 2013). Amongst the resulting candidate strategies, the half-Kelly strategy have been the one most often adopted by practitioners. Kadane (2011) proved that the half-Kelly strategy does not perfectly correspond to optimization of any utility function, but argued that there must exist a constant that is able to account for parameter uncertainty so that logarithmic utility is still optimized. Baker and McHale (2013) found that this optimal shrinkage parameter depends upon the uncertainty related to estimation of both μ and σ[1].

[1] See Appendix D for details on the estimation of μ and σ

We argue that this uncertainty can be linked to what state the market is in, and propose a novel, time-varying approach towards reducing the optimal fraction. The level of uncertainty in parameter estimates is likely to be higher during periods of financial turmoil as compared to periods of relative financial stability. The reasoning is that the deviation from estimated drift $\hat{\mu}$ is expected to be larger during times of significant price declines, where the actual observed drift, μ, is negative. Furthermore, the uncertainty in the estimated volatility, $\hat{\sigma}_t$, is likely to be high in the periods leading up to price declines. This is because, although closely situated time-wise, the volatility is high during the declines (Herwartz and Kholodilin, 2014) and, in general, low during the periods leading up to them. Naturally, it is not possible to determine the actual current market state (defined ex-post). However, the multinomial model continuously generates a probability of being in each of the states, and may be used to reduce the Kelly fraction at times where the parameter uncertainty is assumed to be high. This translates into using the probability of being in a tranquil state, denoted π_{0t}, to obtain the following time-varying, fractional Kelly.

$$f_t^* = \hat{\pi}_{0t} \cdot \frac{\hat{\mu} - r_f}{\hat{\sigma}^2} \tag{6.8}$$

In the original solution to the portfolio problem, Merton showed that the optimal fraction to invest in the risky asset is independent of the level of consumption. By assuming a constant relative risk aversion (CRRA) utility function, the closed form solution to the allocation problem takes the form illustrated in Equation 6.9. The optimal fraction of total investments placed in the risky asset is denoted by f^*, and is inversely proportional to investor risk aversion represented by γ.

$$f^* = \frac{\mu - r}{\sigma^2 \gamma} \tag{6.9}$$

Note that the reduction factor π_{0t} is equivalent to the inverse of the risk aversion term in the Merton fraction presented in Equation 6.9, i.e. $\hat{\pi}_{0t} = 1/\gamma$. Thus, an increase in the risk of financial turmoil is analogous to an increase in investor risk aversion.

We further propose the use of time-varying volatility instead of sample variance, since the optimal allocation problem is solved on a daily basis. This time-varying volatility is assumed to better reflect the current market situation and thus be a better estimator for the actual volatility. Considering that periods of strong price decline go along with considerable increases in stock market volatility (Herwartz and Kholodilin, 2014), a benefit of using time-varying volatility is that it may help reduce the fraction invested during times of financial turmoil. In times of perceived stability, however, using a time-varying volatility may take on excessive risk compared to using a constant sample variance, thus increasing susceptibility to sudden price declines. Adding the restriction of no borrowing, we obtain the final market exposure:

$$f_t^{kel} = \min \left\{ \frac{\hat{\pi}_{0t} \cdot (\hat{\mu} - r_f)}{\hat{\sigma_t}^2}, 1 \right\} \tag{6.10}$$

An important notion is that the Kelly criterion is derived from a situation in which a single bet is repeated an infinite number of times with the same characteristics in terms of odds and probability distribution (Thorp, 2011). Applied to portfolio allocation problems, it leads to an assumption that the market conditions faced when making the investment decision, are faced an infinite, or sufficiently large, number of times. Implicit is an assumption that given these conditions, the market will react identically each time such that the first and the second order moment of the return distribution is similar at each occurrence Thorp (2011).

Given that the aforementioned assumptions hold, the Kelly criterion has been proven to outperform any other fundamentally different strategy for investors with a long investment horizon Kelly Jr (2011). Breiman et al. (1961) even proved that as the number of sequences grows infinitely large, the accumulated wealth from investing according to the Kelly criterion grows infinitely larger then that of any other fundamentally different strategy.

6.4 Trading Assumptions

The following trading assumptions seek to make the trading as realistic as possible. This involve including market frictions such as brokerage commission and bid-ask spread. Several papers, such as Kornprobst (2017), exclude these market frictions.

OSEBX

In this study, the OSEBX is treated as an Exchange Traded Fund (ETF), which enables trading in real-time. However, to our knowledge, no such ETF currently exists. ETFs on the OBX Total Return Index, such as OBXEDNB or OBXEXACT, are possible replacements.

Liquidity

The ETF is assumed to be liquid; a seller is present for the OSEBX value, while a buyer is present for the same value less the bid-ask spread.

Bid-ask spread

The OSEBX level is assumed to be buy prices, and the bid-ask spread therefore only incur when selling. The spread is assumed to be 0.1 % of the trade value, in line with the historical percentage difference between the OBX BID and OBX ASK indices in the period 2014-2019.

Brokerage commission

Brokerage commissions are assumed to be 0.05 % of the trade value (no fixed costs), and incur for all transactions.

Rebalancing constraint

In a dynamic portfolio optimization setting, the importance of continuously updating market exposure according to model specifications must be balanced against the transaction costs incurred when rebalancing. To avoid excessive trading and inflated transaction costs, no transaction is made unless the deviation between the current and desired portfolio weighting is greater than a given threshold. A threshold of 10 p.p. is chosen, meaning that if the current weighting stands at 80 %, rebalancing only occur if the desired weighting is either below 70 % or above 90 %.

Risk-free rate

Uninvested cash receive risk-free rate of return in accordance with the 3-Month Norwegian Government Bond rate. The short duration of this bond conforms well with the investment horizon inherent in the trading algorithms presented in the following sections.

Trading dynamics

The calculations governing the desired market exposure any given day is based on the indicator values and closing prices from the previous day. Any transactions executed to rebalance the portfolio is also assumed to be at the closing price. The implicit assumptions is that an investor have all information needed to dictate desired exposure for tomorrow available at the time of closing, and is able to buy at the exact closing price. The latter is a reasonable approximation if the discrepancy between official closing prices and near-closing prices is small.

Chapter 7

In-Sample Performance and Model Selection

The following chapter first presents an initial analysis of the in-sample performance of a predefined Reference model containing all indicators. The analysis aims to investigate the model's ability to predict the dependent variable on which it is estimated, and presents the estimated probabilities of being in the different states. However, since the ability to classify the dependent variable in the training set is not representative for the model's ability to generalize onto independent datasets, we make no inference on the relative performance of different variable compositions in this part.

In order to determine the desired variable composition and trading strategy, we conduct preliminary out-of-sample testing in the validation set. The performance of each model is evaluated on the basis of statistical metrics as well as direct measures of trading performance through realized return and Sharpe ratio. The chapter is concluded with the selection of a final variable composition that will constitute the candidate strategy for further out-of-sample testing.

7.1 In-Sample Performance

The purpose of the training set is to establish a link between the market states and the indicator levels; the parameters are optimized to minimize estimation errors. Since the final candidate model is yet to be determined, we present the probabilities of being in each state following the estimation of the Reference model. This model contain all indicators, and the 250 days rolling average is chosen for all transformed variables.

As discussed in Chapter 6, the probability of a tranquil state will function as the basis of our trading strategies. The rationale is that given an impending correction or crisis, it is not necessarily important whether it is the probability of correction or crisis that increase, as long as the probability of being in tranquil state is reduced. Figure 7.1 shows the probability of being in this tranquil state, along with the actual definition marked in orange.

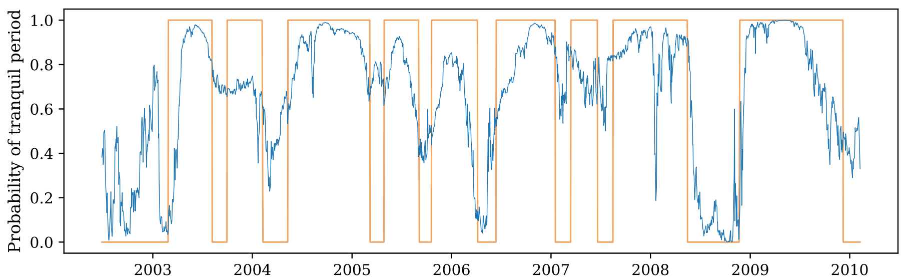
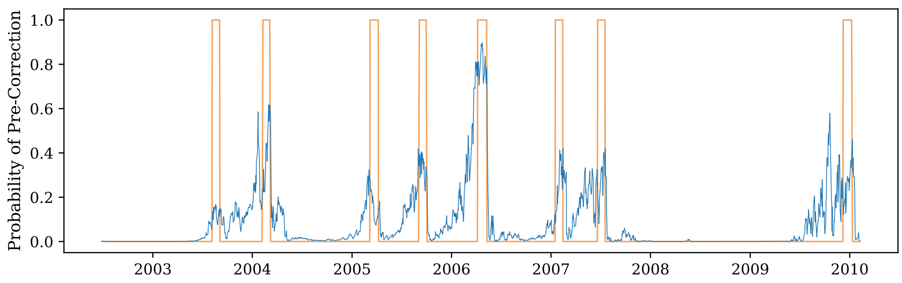


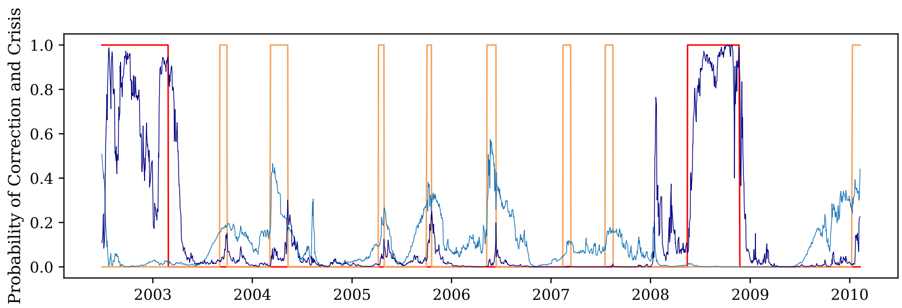
Figure 7.1: Probability of tranquil in the training set for Reference model

Seemingly, the indicators are able to explain some of the movement in the dependent variable. The most apparent examples are the correction in 2006 and the financial crisis in 2008, where the probability is low, combined with a relatively quick recovery after the turmoil has passed.

The probability of being in a pre-correction state is shown in Figure 7.2a, while corrections and crises are shown in Figure 7.2b. There seem to be a good fit between the pre-correction period and the predictions, which are low during other states, and clearly increasing when approaching a correction. The probability is also sharply decreasing after a correction start, indicating an ability to recognize the characteristics of sharp price declines. In the case of crisis and correction periods, we observe that the probability of crises is more distinct. This may originate from a notion that the indicators take their outermost values under the extreme movements reflected in financial crises, making their levels easier to recognize.



(a) Probability of pre-correction



(b) Probability of corrections (light blue) and crises(dark blue)

Figure 7.2: Probabilities of pre-correction, correction and crisis in the training set for Reference model

7.2 Model Selection

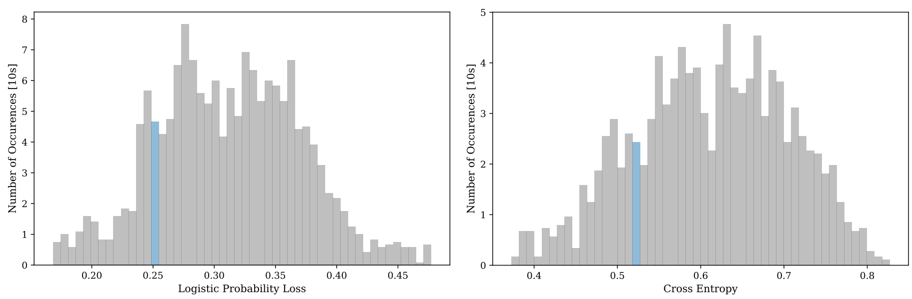
The purpose of the validation set is to identify the model which is expected to generalize best on an independent dataset. To this extent, we proceed with comparing the performance of the fitted models using a combination of statistic- and financial metrics.

Statistical Metrics: LPS and Cross Entropy

In order to determine the variable composition of the final model specification, we first utilize the statistical measures presented in Section 5.4. Table 7.1 presents the performance of the Reference model and the three best performing models (lowest values) according to the LPS criterion, along with their Cross Entropy scores. Figures 7.3a and 7.3b show the distribution of LPS and Cross Entropy scores, respectively, in the validation set across all models.

Model	LPS	Cross Entropy
PB250 PE120 EQ250 SX250 SY250 VIX	0.168	0.389
PB250 PE120 EQ120 SX250 SY250 VIX	0.169	0.390
PB250 PE120 SX250 SY250 VIX	0.169	0.390
PB250 PE250 EQ250 SX250 SY250 CCI VIX VSI	0.253	0.526
No. observations	1032	

Table 7.1: LPS and Cross Entropy of selected models in validation set



(a) LPS

(b) Cross Entropy

Figure 7.3: Distribution of LPS and Cross Entropy across all models in validation set, with Reference model (blue)

Judging from the LPS criterion, the best performing model consist of all variables, with the exception of CCI and VIS. In order to investigate this further, Figure 7.4 shows the probability of being in a tranquil state for the best performing model, along with the base case with all variables included.

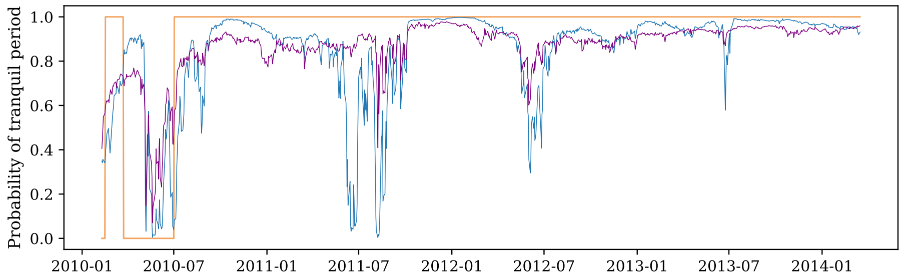


Figure 7.4: Probability of tranquil in validation set for best LPS (purple) and Reference (blue)

Apparently, the base model wrongfully suggests high risk in several time periods in which there were no crisis or correction. However, by inspecting the relationship between these probabilities and the actual price movements, the impression is altered. As evident from Figure 7.5, there was indeed a clear price decline around the periods of high predicted risk. While the model correctly lower the probability, it is punished since no correction or crisis is defined.

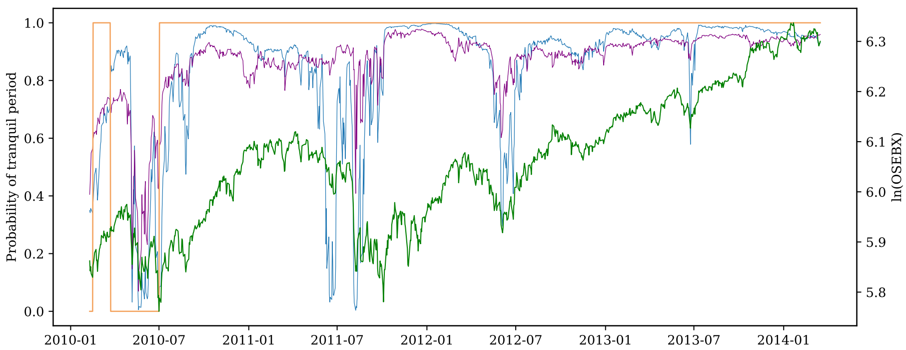


Figure 7.5: Tranquil probability for best LPS (purple) and Reference (blue). OSEBX (green)

This example represents an inherent issue in utilizing the LPS metric in this context. A significant loss of information is incurred when transferring the price movements of the OSEBX into a limited dependent variable, which in turn infect the analyses made on the basis of this simplification.

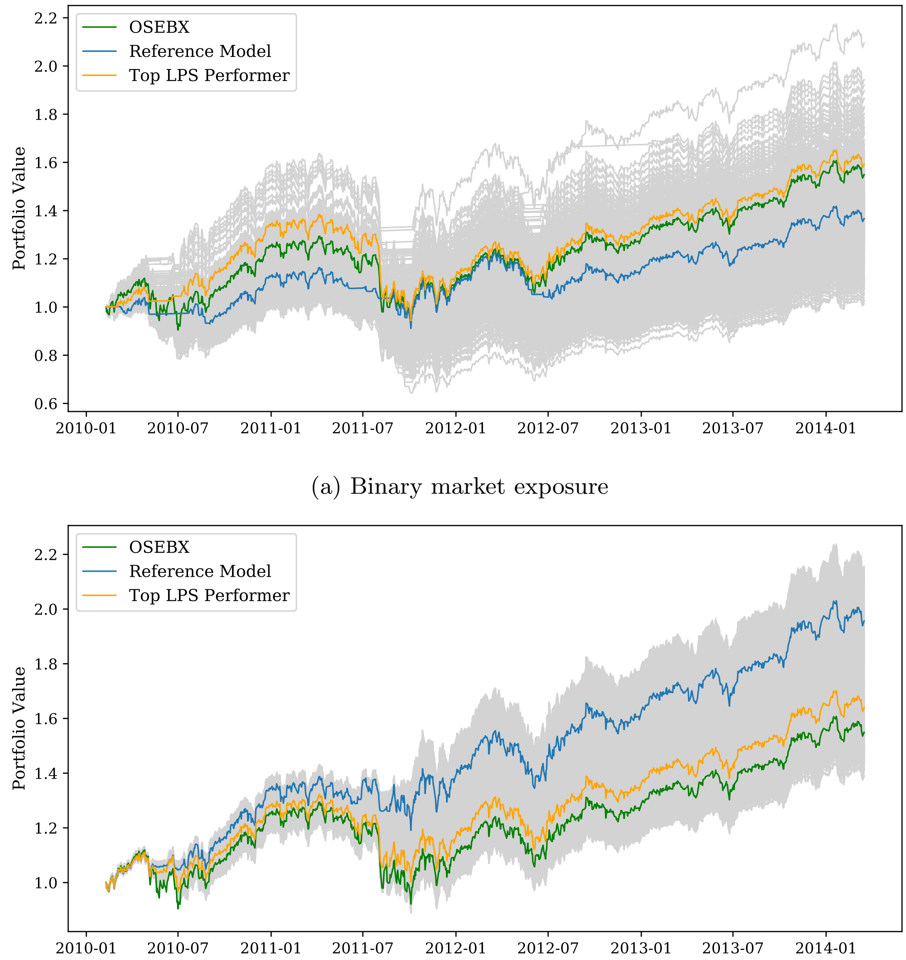
Financial Metrics

The benefit of using realized returns and the Sharpe ratio is that the metrics are independent of the definition of the dependent variable. Information loss incurred when transforming the price movements into binary indicators of different states will therefore not impact the measured performance of our model. Furthermore, while the statistical metrics measures the divergence from the dependent variable symmetrically, returns will better reflect the financial consequences of wrongful predictions.

One drawback of the financial metric approach, however, is the reliance on a trading strategy. If chosen inappropriately, the strategy may fail to translate the probabilities into meaningful market exposures, and hence impair the performance of an otherwise functioning EWS. On the other hand, an exceptional strategy may yield favorable results regardless of the input probabilities, thereby undermining the impact of the probabilities themselves.

Meanwhile, the purpose of the validation period is to assess any model's ability to generate risk-adjusted return in the test period. Given that the selected candidate strategy is able to display a satisfactory level of generalization ability, the risk-adjusted return in the validation period should theoretically be a valuable measure of performance in the subsequent test period (Friedman et al., 2001). Recognizing these considerations, we argue that risk adjusted return is a more representative metric of expected performance in the final test set. Thus, the Sharpe Ratio will enact as the principal determinant in the proceeding model selection. Both the LPS and Cross Entropy scores will, however, be used to support the decision.

After selecting the desired performance metric, the trading strategy must be determined. Figure 7.6a displays the returns of OSEBX compared to the Reference and best LPS model using the binary market exposure. Figure 7.6b shows the same, only using gradual market exposure.



(a) Binary market exposure

(b) Gradual market exposure

Figure 7.6: Portfolio values for all models in validation set, with best LPS, Reference and OSEBX highlighted

At first glance, both strategies seem to generate returns in excess of the OSEBX. We note that the base model underperforms the market in terms of return when using the binary strategy, while overperforms in the Kelly framework. The model with the highest LPS score marginally outperforms the market in both cases. In order to make any inferences on the risk-adjusted performance of the models, Tables 7.2 and 7.3 present the five models with the highest Sharpe Ratio for binary and gradual exposure respectively. Figure 7.7 shows the distribution of realized Sharpe ratios in the validation set across all models.

Model	Total Return[%]	Annual Return[%]	Vol[%]	Sharpe Ratio
PB250 EQ250 SY250 VIX	110	20	16	1.10
PE120 EQ250 SX120 SY250 CCI VIX	90	17	15	1.01
EQ250 SX120 SY250 CCI VIX	86	16	15	0.95
PE120 SX120 SY250 CCI VIX	83	16	15	0.94
PB250 PE250 EQ250 SX250 SY250 CCI VIX VSI	37	8	16	0.38
Market Portfolio	55	11	20	0.47

Table 7.2: Performance of top four models and Reference model with binary exposure in validation set

Model	Total Return[%]	Annual Return[%]	Vol[%]	Sharpe Ratio
PB250 PE250 EQ250 SY120 CCI VIX	113	20	16	1.19
PB250 PE250 CCI VIX	111	20	15	1.18
PB250 PE250 EQ120 CCI VIX	112	20	16	1.18
PB120 PE120 EQ120 SX120 SY120 CCI VIX	116	21	16	1.18
PB250 PE250 EQ250 SX250 SY250 CCI VIX VSI	96	18	17	0.94
Market Portfolio	55	11	20	0.47

Table 7.3: Performance of top four models and Reference model with gradual exposure in validation set

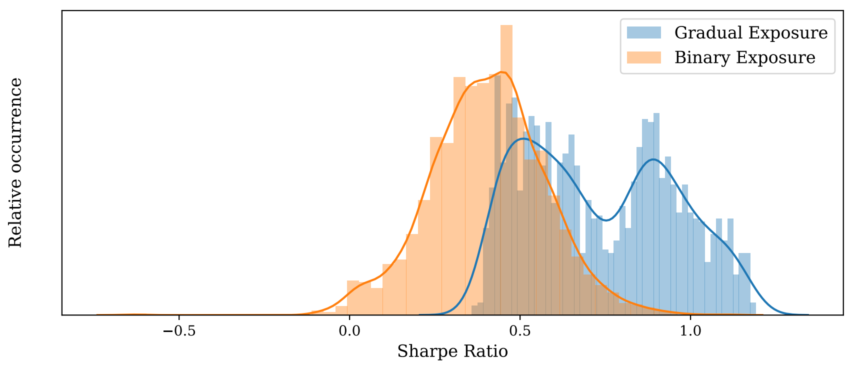


Figure 7.7: Distribution of Sharpe Ratios for binary and gradual exposure in validation

In general, the strategy based on the Kelly criterion seems to outperform the binary marked exposure. A buy-and-hold strategy in the OSEBX yields a Sharp ratio of 0.47, which approximately equals the mean Sharpe ratio of the models following a binary market exposure. Meanwhile, the majority of models following the Kelly strategy outperforms the OSEBX, with the top performing model reaching a Sharpe ratio of 1.19. Consequently, we establish the Kelly based strategy as the preferred choice of trading strategy in further analysis.

When choosing the desired variable composition, we turn to the models exhibiting the highest Sharpe ratio, as displayed in Figure 7.2. The top performing model consist of all variables with the exception of SX and VSI. The time intervals used in the rolling average to create the transformed variables are all 1 year, with the exception of a half year for the SY variable. With most variables included, and all groups of indicators are represented, the model seems reasonable. We therefore proceed to investigating the estimated coefficients of a Candidate model consisting of PB250, PE250, EQ250, SY120, CCI and VIX, with the use of the Kelly based trading algorithm.

Interpretation of Estimated Coefficients

Before proceeding with out-of-sample testing, we investigate the effects of each indicator on the different probabilities. Table 7.4 show the estimated coefficients for the current Candidate model.

	β_{i0}	β_{i1}	β_{i2}	β_{i3}
PB	-7.30	1.69	3.51	2.10
PE	0.21	1.49	0.38	-2.07
EQ	1.08	-0.49	0.27	-0.86
SY	-4.53	0.55	-0.28	3.71
CCI	0.78	0.02	-0.89	0.09
VIX	0.02	-0.20	-0.02	0.20
Intercept	2.07	1.75	0.93	-4.74
No. observations	1915			

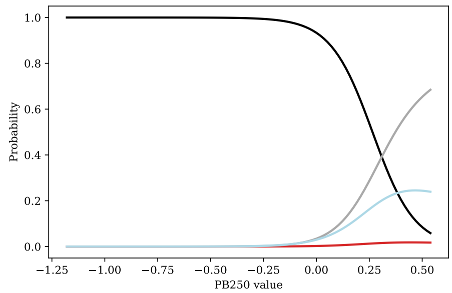
Table 7.4: Coefficients of top performing model in validation set

The logistic regression model is most naturally interpreted in terms of the effects on the bilateral relationship between pairs of categories of the dependent variable. This is easily illustrated from the linear predictor function presented in Section 5.2,

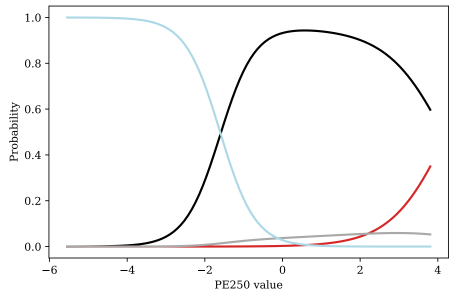
$$\eta_{it} = \ln\left(\frac{\pi_{it}}{\pi_{0t}}\right) = \alpha_i + \boldsymbol{\beta}_i \boldsymbol{x}_t \ , \quad \forall\ i \in \{1, 2, 3\}$$

In a binomial logistic regression model, there is only one such pair, and the impact of each variable is reasonably easy to interpret given a set of coefficients. When considering a multinomial model, however, where the dependent variable can take more than two values, the interpretation is not so straightforward. As the numbers of categories increase, the interpretation becomes progressively challenging. In the case of a four-state dependent variable, no sole coefficient specific to an indicator can be impartially interpreted without the other three. Furthermore, the relationship is non-monotonous, meaning its impact may vary depending on the values of other variables.

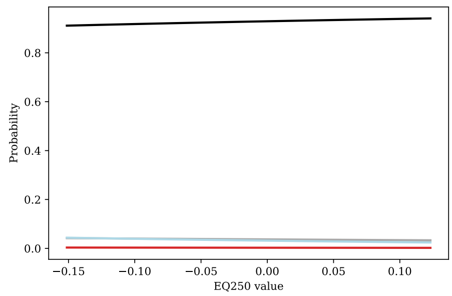
In an attempt to assess the impact of the indicators on the probability of crises and correction, we employ a graphical interpretation method introduced by Long (1997). The method seeks to illustrate the dynamics between the effects by plotting the different probabilities when changing one variable and holding all others constant. The results are presented in Figure 7.8.



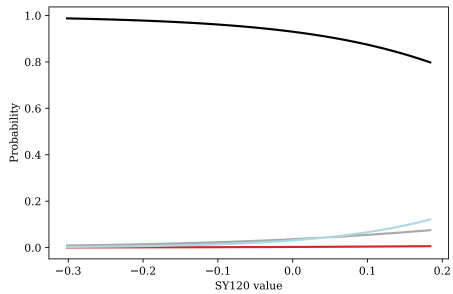
(a) PB250



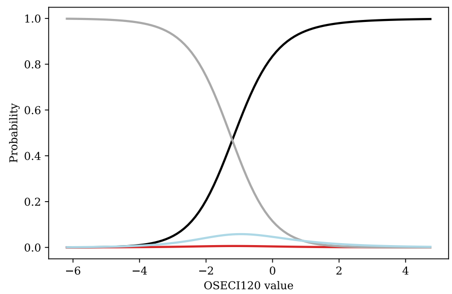
(b) PE250



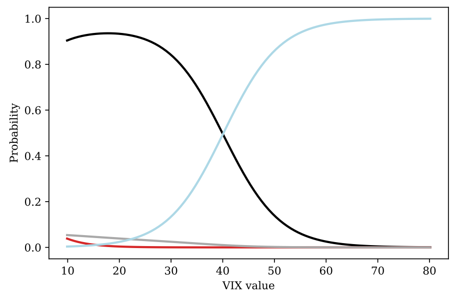
(c) EQ250



(d) SY120



(e) CCI



(f) VIX

Figure 7.8: Marginal effects of indicators on the probability of being in a tranquil (black), pre-correction (red), correction (grey) or crisis (blue) period

An important notion is that the dynamics of the relationships illustrated in Figure 7.8 rely on isolated changes when all other variables are held constant. Thus, the presence of any significant correlation between the indicators would distort the discussion of these relationships. As an example, Figure 7.8e seemingly convey that an increase in the CCI implies an increase in the probability of being in a tranquil state. However, noting that the CCI is positively correlated with both the PB and PE ratios and negatively correlated with the VIX, its impact is less clear. Evident from Figures 7.8a, 7.8b and 7.8f, the changes in these values implied by the correlations may in fact contradict the impact of the CCI alone. We note that while the plots may convey the apparent impact of the various indicators, any direct inference may be misleading.

However, while interpretation of the effects is hard due to indicator correlation, two variables display an apparent lack of *any* impact on the resulting probabilities. Figures 7.8c and 7.8d show that despite shifting either the EQ250 or SY120 from their historical minimum to the historical maximum, the impact on any probability is severely limited. Furthermore, the impact of the EQ250 and SY120 seem to be rather insignificant from a trading perspective. In fact, Table 7.7 show that removing these two variables only inflict a marginal decrease in the Sharpe ratio, from 1.19 to 1.18.

Moreover, the potential multicollinearity issues discussed in Section 5.3 may lead to inefficient and unreliable coefficients. The common approach to mitigate the issue is to drop one or more of the variables, thereby reducing the collinearity (Kleinbaum et al., 2002). Recognizing the limited impact of the EQ250 and SY120 variables on the Sharpe ratio obtained in trading results, we argue that both should be omitted from further analysis. Thus, the final variable composition consists of the variables PB250, PE250, CCI and VIX. Figure 7.9 shows the returns of this combination in the valuation period, along with the Reference model and OSEBX. The results are summarized in Table 7.5, and the model coefficients are shown in Table 7.6.

Plot of Final Candidate Strategy

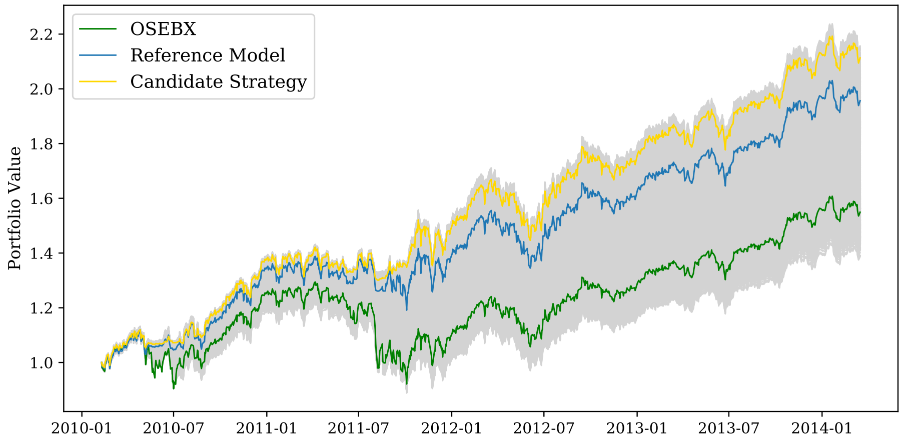


Figure 7.9: Portfolio values of all models in validation set, with OSEBX, Reference model and Candidate strategy highlighted

Summary of Performance

Model	Total Return[%]	Annual Return[%]	Vol[%]	Sharpe Ratio
PB250 PE250 CCI VIX	111	20	15	1.18
Market portfolio	55	11	20	0.47

Table 7.5: Performance of Candidate strategy vs. market in validation set

Coefficients

	β_{i0}	β_{i1}	β_{i2}	β_{i3}
PB	-7.86	1.86	3.57	2.42
PE	0.24	1.46	0.36	-2.07
CCI	0.78	0.02	-0.88	0.09
VIX	0.02	-0.20	-0.01	0.19
Intercept	2.00	1.56	0.79	-4.35
No. observations	1915			

Table 7.6: Coefficients of Candidate strategy

Chapter 8

Out-of-Sample Performance

This chapter presents and discusses the final out-of sample results in the period march 2014 - march 2019. First, the performance of the Candidate strategy chosen in Chapter 7 is compared to the Reference model and a buy-and-hold strategy in the OSEBX, while a series of random portfolios is utilized in the analysis of the statistical significance of our results. These tests will provide the foundation for the inference made on the validity of the market efficiency hypothesis on the Oslo Stock Exchange.

A second part presents a more detailed discussion of the estimated probabilities and resulting trading performance. Here, we aim to investigate the apparent strengths and weaknesses of our framework, in order to better understand its limitations. This includes a review of the output probabilities from the multinomial logistic regression model, along with a discussion of the effect of the Kelly criterion on the ultimate market exposure fractions. Finally, the chapter carry out an analysis on the effects of individual indicators.

8.1 Trading Results

As a preliminary performance assessment, we investigate the returns and Sharpe ratios of the different strategies. Figure 8.1 illustrate the paths of both the Candidate strategy, the reference model and the market, while Table 8.1 shows the corresponding Sharpe ratios.

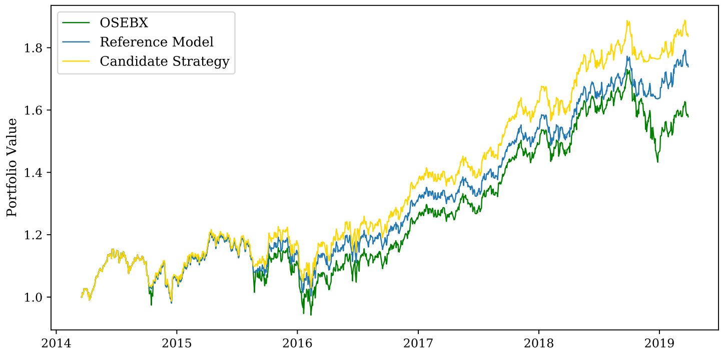


Figure 8.1: Portfolio value of OSEBX, Reference model and Candidate strategy in test set

Model	Total Return[%]	Annual Return[%]	Vol[%]	Sharpe Ratio
Candidate	84	13	14	0.86
Candidate ex. market frictions	87	13	14	0.89
Reference Model	74	12	14	0.79
Market portfolio	58	10	16	0.55

Table 8.1: Performance of Candidate strategy with- and ex. market frictions

Evidently, the Candidate strategy outperforms both the reference model and the OSEBX. Although somewhat late, it correctly reduces exposure during the downturns in 2016 and 2018, while also displaying ability to maintain full exposure in times of significant growth. Although only one of the significant declines in the test set is captured by our definition, the Candidate strategy is able to outperform the market by 26 p.p. in terms of total return. Excluding market frictions, the Sharpe Ratio increase from 0.86 to 0.89. The algorithm executes a total number of 76 trades with transaction costs amounting to 2.4% of initial wealth.

8.2 Assessment of Significance

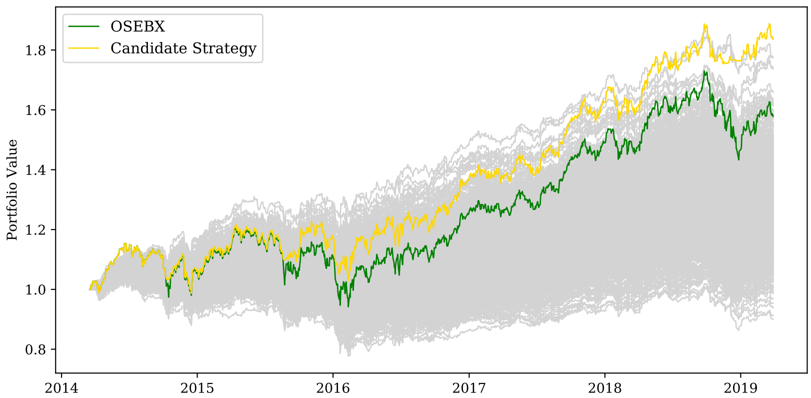
In order to make any inference on the statistical significance of our results, we construct a series of random portfolios using Monte Carlo simulation. The basic idea is to assess whether any eventual success of the Candidate strategy is simply a matter of luck, or if the predictive power of the framework is significant.

Construction of the random portfolios will follow three different approaches. The first involve creating fully arbitrary portfolio weights which determine the daily market exposure. Each portfolio is updated on a daily basis, and earn the appropriate return. While this procedure enables complete randomness in calculating fractions, the average exposure will not necessarily resemble that of the Candidate strategy.

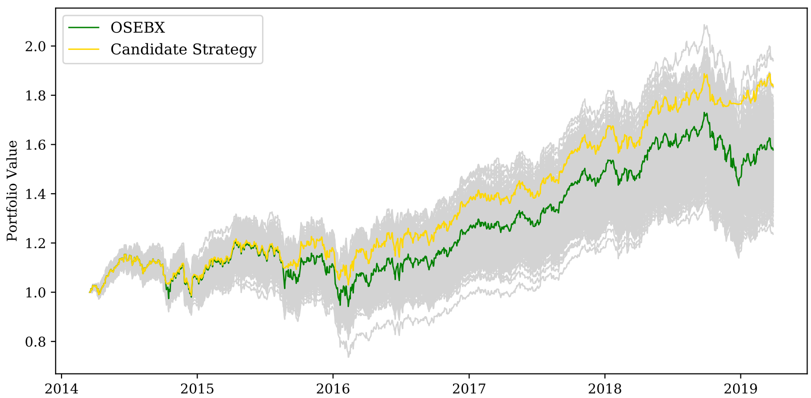
Similar to the *Random Same Proportion* (RSP) portfolios of Douady and Kornprobst (2018), we create random permutations from the list of exposure fractions given by the Candidate strategy. With the fractions drawn without replacement, this procedure ensures an average exposure equaling that of the Candidate strategy. As with the fully arbitrary approach, each portfolio is updated daily to a new, quasi-arbitrary exposure and earn the appropriate return.

In order to isolate the contribution of our multinomial logistic regression model, we create a third set of random portfolios. This procedure also involves random permutations, but this time on the list of probabilities given from the EWS. The permuted probabilities will thus have the same mean as the probabilities of the Candidate strategy, and function as the input when calculating the Kelly fraction. This way, we aim to distinguish the effect of the EWS probabilities from the contribution of the Kelly criterion.

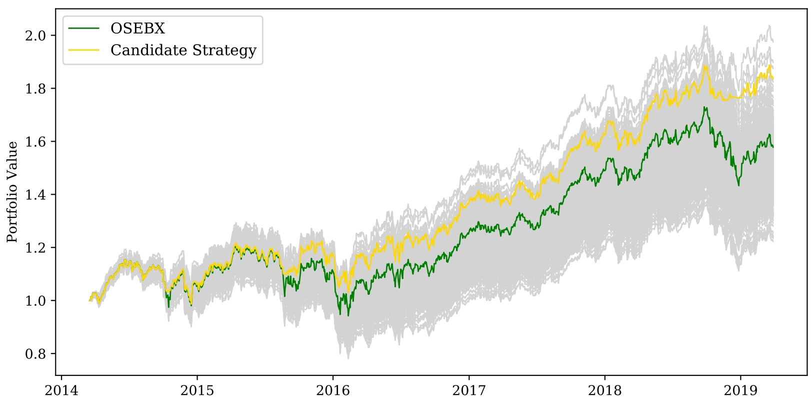
Since randomly drawn market exposures may incur frequent rebalancing and an artificially high number of transactions, we omit transaction costs in the simulation of random portfolios. Figure 8.2 shows the returns of the Candidate strategy compared to all three series of random portfolios, each with 1 000 simulations.



(a) Fully arbitrary exposure



(b) Random same proportion exposure



(c) Random same proportion probability

Figure 8.2: Portfolio value of randomly generated, OSEBX and Candidate strategy in test set

The dispersion of the fully arbitrary random portfolios in figure 8.2a is clearly larger than that of the same proportion portfolios in figures 8.2b and 8.2c. Since the latter are governed by permutations of given fractions or probabilities, these will more similarly resemble the risk characteristics of the Candidate strategy averaging 94 % market exposure.

The Candidate model appear to outperform all fully arbitrary portfolios and the majority of the same proportion portfolios. Since the each of former portfolios have an expected average market exposure of 50 %, the predominance of the Candidate strategy in terms of absolute return may largely be associated with a higher level of risk. For the same proportion portfolios, however, the Candidate strategy seem to display a performance that exceed that of random portfolios with similar risk characteristics.

In order to investigate the risk-adjusted performance of the different strategies, we evaluate the distributions of the Sharpe ratios in comparison to the different random portfolios. Figure 8.3 shows the distribution of Sharpe ratios for 10 000 realizations of all three series of random portfolios.

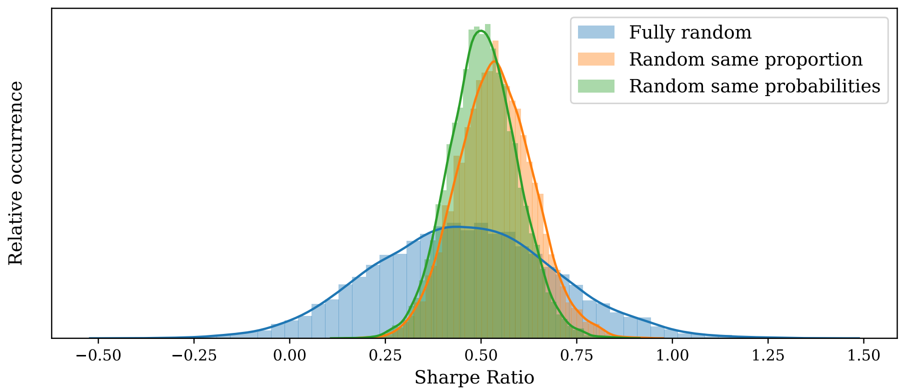


Figure 8.3: Distribution of Sharpe ratios for fully random, random same proportion and random same probabilities

The Sharpe ratio of the Candidate amounts to 0.86. When compared to the distributions of the fully random-, same fraction- and same probability portfolios, this performance correspond to the 95th, 99.9th and 99.99th percentiles, respectively.

The Candidate strategy outperforms 99.9 % of the random same proportion portfolios. Since market exposures for both the Candidate and the random portfolios are governed by the Kelly criterion, this indicate that the exposure fractions resulting from the Candidate strategy have a significant impact on resulting performance relative to other portfolios with similar risk characteristics.

Compared to the same probability portfolios, the Candidate strategy outperforms 99.99 % of the random portfolios. This means that the market exposure resulting from the Kelly criterion is largely dependent on the input probabilities, and we thus conclude that these probabilities carry relevant information for timing the market and in turn achieve a high Sharpe ratio.

Comparing our results to the fully random portfolios, we conclude that the market has been inefficient of semi-strong form at a 95 % significance level. This prove the potency of our approach, and similar to Kornprobst (2017), we show that winning trading strategies can be created in the field of large financial downturns - even after the inclusion of transaction costs.

8.3 Discussion

Following the statistical significance of our results, the output probabilities of the multinomial regression model seem to carry explanatory content for the price movement of the OSEBX. Figure 8.4 shows the estimated probabilities of pre-correction and tranquil periods in the test set.

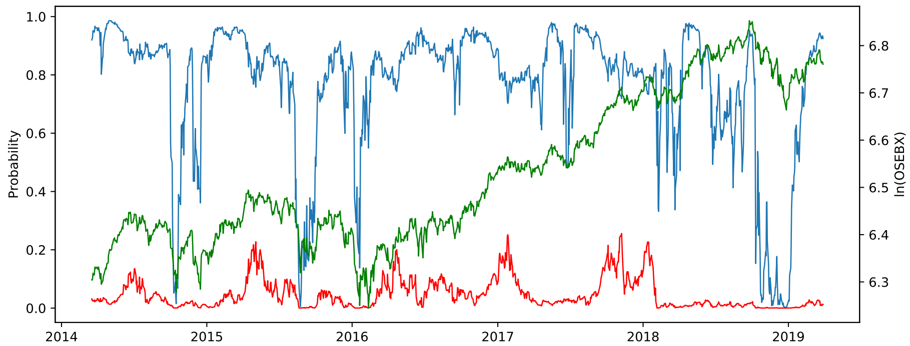


Figure 8.4: Probability of tranquil (blue) and pre-correction (red) in test set. OSEBX (green)

Evidently, low probability of being in a tranquil period seems correlated with periods of price decline. Meanwhile, the explanatory power of the probability of pre-correction periods looks rather limited, never exceeding 25 %. This is supported by the graphical illustrations of the indicators' marginal effects, as shown in Section 7.2. With the exception of PE250, the indicator's impact on the probability of pre-correction seems limited, and the ability to reduce the market exposure before a price decline is therefore likely to be equally limited.

If not able to identify pre-correction, the model is seemingly able to promptly identify price declines. The effect on the probability of tranquil is substantial; during the correction in 2018, the probability of being in a tranquil state is only 0.2 %. One explanation of why the model appears to better recognize an ongoing correction than it's build up can be found in possible reversed causality. Since the underlying price movement governing the dependent variable is also found in the mathematical formulation of several indicators, it is evident that the levels of the indicators during large price declines (especially when compared to a recent mean) are largely affected by the price decline itself.

Figure 8.5 shows the investment fraction of our Candidate strategy throughout the test period. It is evident that the aforementioned discrepancy in the models ability to recognize the different states materializes in the observed market exposure. The impact of the probabilities of crisis and correction is clearly recognizable in the fractions invested. They seem to correctly contribute to a reduction of market exposure during periods of market turmoil, such as in 2014, 2015, beginning of 2016 and the correction in 2018. These reductions make the model achieve excess return compared to the market.

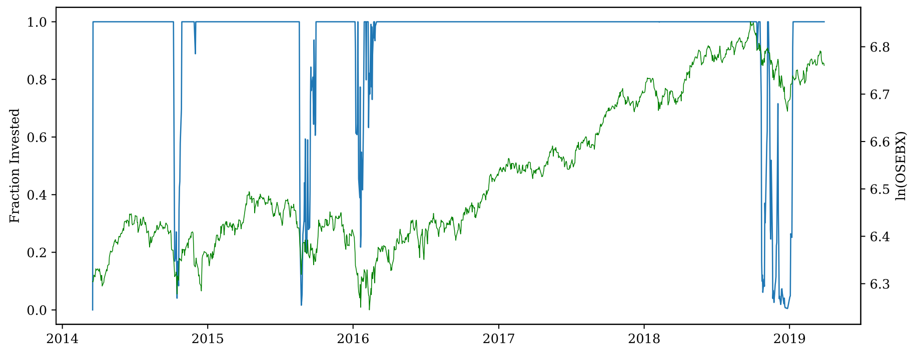


Figure 8.5: Market exposure of Candidate (blue) compared to ln(OSEBX) (green) in test set

When comparing Figures 8.4 and 8.5, the footprint of pre-correction is undetectable in the resulting market exposure. The impact of pre-correction on the probability of tranquil is not large enough to reduce the market exposure below one, making the model enter the price declines fully invested. A possible extension of the trading rule is to consider the probability of pre-correction isolated to mitigate this risk.

For the Candidate strategy to maintain full exposure and ignore the effects of increasing probability of pre-correction, the fraction implied by the Kelly criterion must be correspondingly high during the given period. In order to investigate these dynamics, Figure 8.6 show the exposure of our Candidate strategy when relaxing the constraint of no gearing.

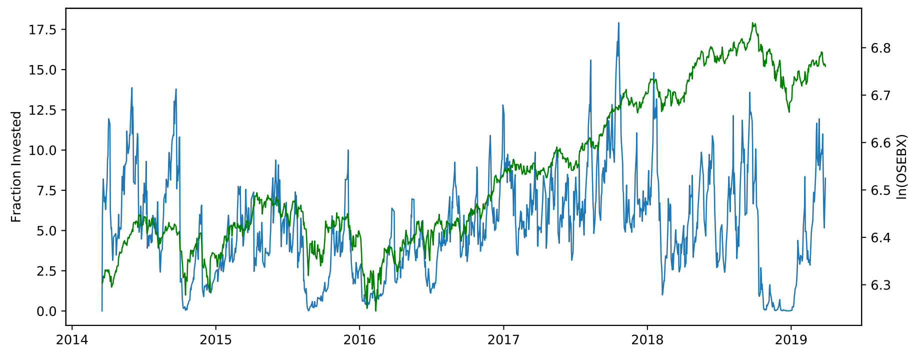


Figure 8.6: Exposure of relaxed Candidate strategy (blue) against ln(OSEBX) (green)

Figure 8.6 shows that the unconstrained strategy generally suggest a very high market exposure. When gearing is allowed, the average exposure increases from $\overline{f^*} = 0.94$ to $\overline{f_g^*} = 5.27$. It is also interesting to note the tendency of high exposure to occur prior to significant price declines. This accentuates the notion that the probability of pre-correction fail to impact the final investment fraction in a satisfactory manner. One possible explanation is that periods of sustained positive growth are characterized by low volatility levels, causing the fraction implied by the Kelly criterion to rise. Consequently, our strategy is not able to pick up the signals of an impending market correction. The Kelly criterion causes our Candidate strategy to maintain full exposure during periods of market growth, but this appear to come at a cost of failing to recognize impending market downturns.

The observed exposure of our Candidate strategy in Figure 8.6 take on abnormally high levels, which is evidently caused by a very high fraction implied by the Kelly criterion. Nekrasov (2014) found that when applying the Kelly criterion in multivariate portfolio optimization, the resulting portfolios tended to be undiversified with large exposures towards a limited number of stocks. Ziemba et al. (2003) also note that fractions derived from the Kelly criterion can be large and that the wealth grows on a bumpy path, with large gains followed by large losses. Nevertheless, the Kelly investor is sure to win if the investment horizon is sufficiently long (Nekrasov, 2014).

Under such inflated fractions, it is important to keep in mind the limiting assumptions of the Kelly criterion in real life applications. Firstly, the same conditions in terms of perceived risk(volatility) and expected gain (excess return) must be met an infinite number of times. Baker and McHale (2013) proved that if this assumption is not met, the criterion will be more risk seeking than that which should be implied by the logarithmic utility function. This appears to be a valid concern with our Candidate strategy, as taking on 17.5 times the market risk simply makes no sense for an investor who is fundamentally risk averse.

Additionally, when modelling the risky assets using the Geometric Brownian Motion, we implicitly assume normally distributed asset returns. This means that every time an investment decision is made, it is made according to the return distribution implied by the constant drift and the time varying volatility. The distribution will consequently change in accordance with volatility, as illustrated in Figure 8.7. This figure show the normal distribution of returns implied by the parameter estimations corresponding to the time increments with the highest-, average- and lowest exposure.

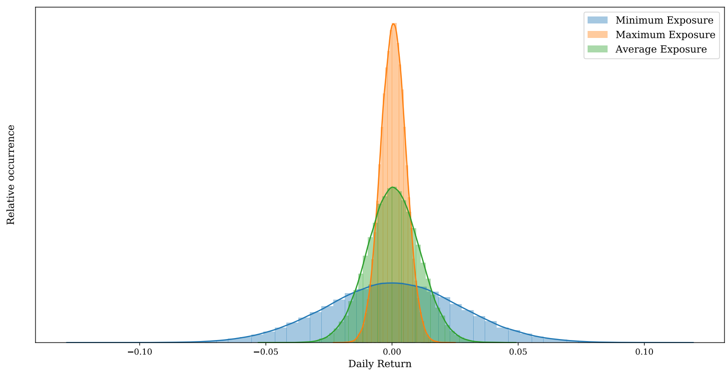


Figure 8.7: Normal distributions with first- and second order moments matching the highest, lowest and average exposure of the relaxed Candidate strategy

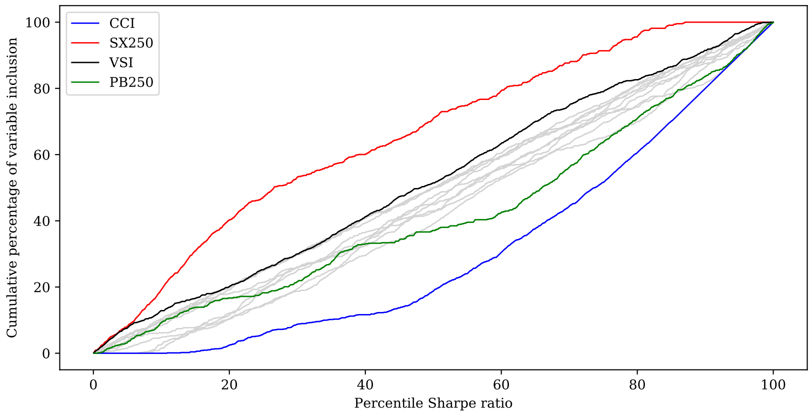
From the orange distribution in Figure 8.7, we observe that daily losses in the excess of 2 % are highly unlikely ($p \approx 0.1\%$). Under such circumstances, the Kelly criterion will perceive the risk as artificially low and consider the pending bet as a favorable one with limited downside. This is obviously not true for real capital markets, where asset returns exhibit skewed distributions with fat tails (Christoffersen, 2011). The resulting trading strategy will be highly susceptible to significant, negative asset price movements which are not captured by the normal distribution.

The final cause of concern is uncertainty in estimation of the out-of-sample return distribution, more specifically μ and σ. Ex-post, we observe that the true drift in the test set was $\mu = 0.0412\%$ notably lower than our estimate of $\hat{\mu} = 0.0557\%$. A consequence of overestimating the true drift is that the fraction implied by the Kelly criterion will always be above its true value. In our Candidate strategy, the probabilities are included to shrink the final exposure. However, if the probabilities fail to reduce the final fraction sufficiently, the consequence could be severe. Baker and McHale (2013) found that failing to shrink the fraction below its true value, as implied by the Kelly criterion, is almost sure to lead to bankruptcy.

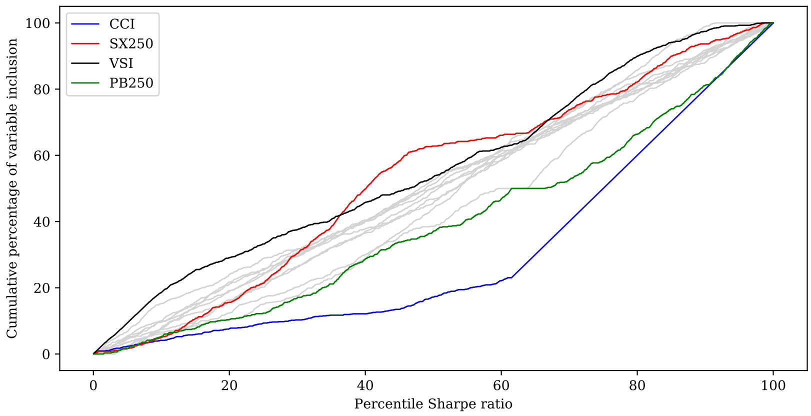
Our results serve to substantiate the notion that with assets that appear favorable, such as the OSEBX, a Kelly investor must exhibit vigilance and sometimes greatly reduce the proposed exposure. Our approach of reducing the exposure according to information contained in fundamental indicators while refusing to take on more risk than that implied by the market has proven to be a winning trading strategy.

Indicator Analysis

As a final analysis, we investigate the effects of individual indicators on the resulting performance. A natural starting point is to investigate all models and their indicator compositions to assess any patterns in terms of which variables that most often appear in the highest and lowest performing models. Figures 8.8a and 8.8b show the cumulative percentage of total variable inclusion plotted against the percentile of Sharpe ratios for all models in the validation and test set, respectively. If the Sharpe ratios are high most of the times an indicator is included, the indicator is assumed to be effective.



(a) Validation set



(b) Test Set

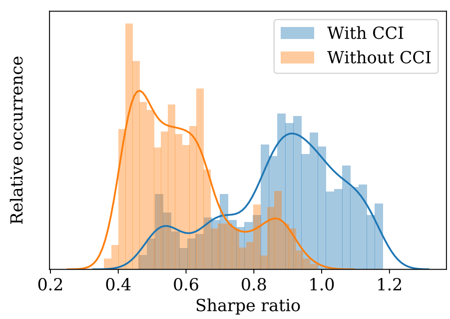
Figure 8.8: Sharpe ratios when the different indicators are included, in validation and test set

It appears that the CCI and PB250 are important determinants of performance, since they are often constituents in the best models. Adversely, SX250 and the VSI seem to mostly be included in the poorer performing models. We proceed with a discussion of the effects of these four variables.

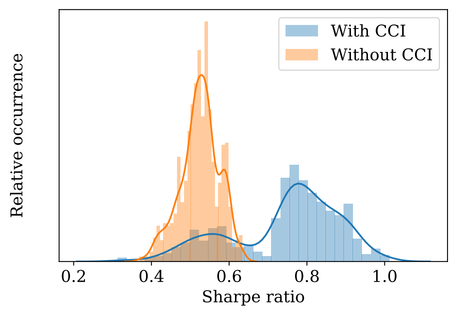
CCI

The CCI (blue line) is seemingly an essential indicator in the validation set. While almost entirely unobserved amongst the models with the 20 % lowest Sharpe ratios, cumulative variable inclusion increases more than that of any other indicator after the 50 % percentile is passed. The importance of the CCI for obtaining a high Sharpe ratio is also clearly shown in the test set; if one chooses a random model where the CCI is included, it is an 80 % chance that the model is amongst the 40 % best.

The importance of CCI is further substantiated by Figures 8.9a and 8.9b, which show the distribution of Sharpe ratios for models with and without CCI in the validation and test set. Models with CCI have, on average, higher Sharpe ratios both in the validation and test set.



(a) Validation set



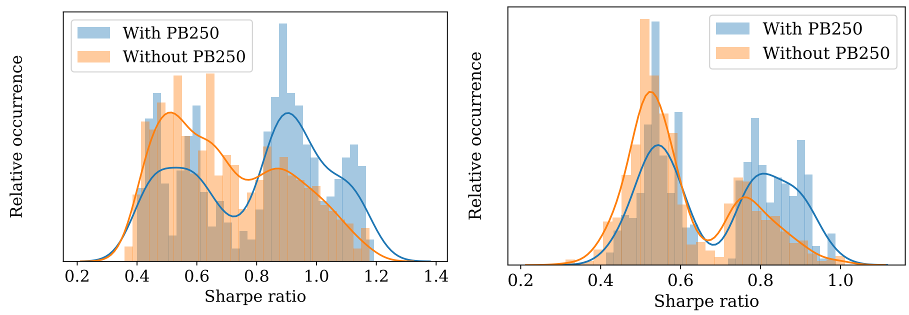
(b) Test set

Figure 8.9: Distribution of Sharpe ratios for models with CCI and without CCI in the validation and test set

Being an indicator solely based on historical prices, this result is somewhat surprising, and may be a sign of momentum effect in OSEBX. The result is difficult to compare with other studies (e.g Patel et al. (2015); Kara et al. (2011)), since they have not analyzed the indicators' marginal effects. However, there is comprehensive documentation of the profitability of momentum strategies (e.g. Jegadeesh and Titman (1993); Jegadeesh and Titman (2001); Jegadeesh and Titman (2011)). Their strategy was to buy the stocks that have performed well in the past, and short sell stocks that have performed poorly. They showed that the anomaly persisted by generating significant positive return over the period from 1965 to 2004.

PB250

Plotted in green in Figure 8.8, PB250 is often included in the models with the highest Sharpe ratios. The importance of PB250 is shown in Figures 8.10a and 8.10b, where the models containing PB250 on average have higher mean. This result is in line with Fu et al. (2019), which conclude that the price-to-book ratio seem to have strong explanatory content for the emergence of price bubbles.



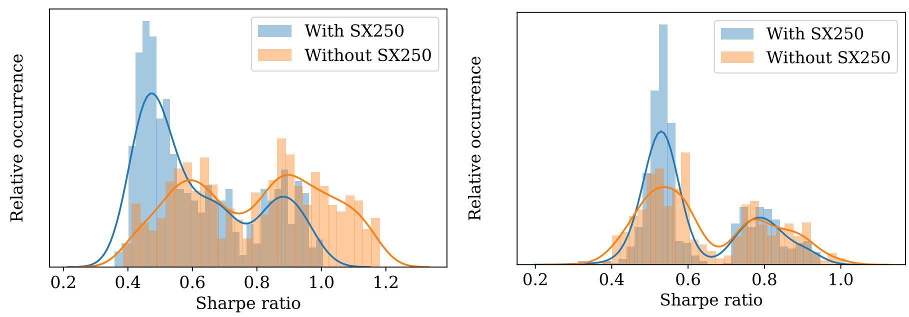
(a) Validation set

(b) Test set

Figure 8.10: Distribution of Sharpe ratios for models with PB250 and without PB250 in validation and test set

SX250

Another observation is that SX250 (red line in Figure 8.8) seem to appear amongst the lowest performing models in the validation set more than any other indicator. Approximately 50 % of the models including SX250 turn out amongst the 30 % lowest in terms of Sharpe ratio. Furthermore, the variable is not included in either of the 15 % best models. Its contribution in the test set, however, is somewhat improved. The distributions of Sharpe ratios for models with and without SX250 are shown in Figures 8.11a and 8.11b, and confirm these observations.



(a) Validation set

(b) Test set

Figure 8.11: Distribution of Sharpe ratios for models with SX250 and without SX250 in validation and test set

The distributions of models including SX250 have a lower mean than that of those that does, both in the test and validation set. Figure 8.12 shows that the link between the indicator levels and the dependent variable is not consistent. As mentioned in Section 4.2, the industry composition of OSEBX and STOXX 600 EUROPE is not the same; Bhojraj and Lee (2002) state that the accuracy of PE is better when comparing to industry peers, which may explain why the indicator turn out inefficient.

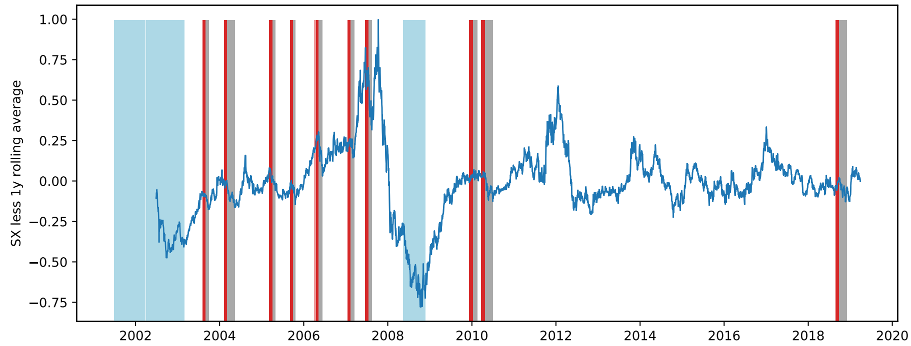
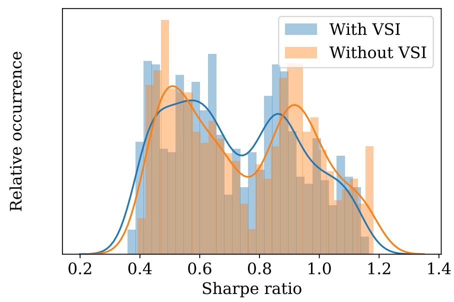


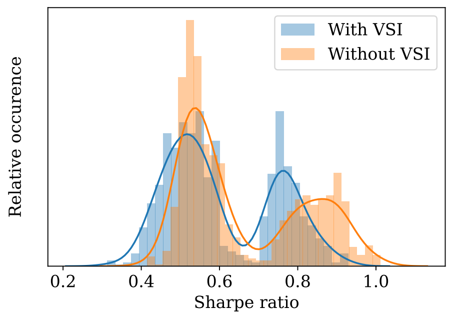
Figure 8.12: PE OSEBX / PE STOXX 600 Europe less its 1 year rolling average

VSI

Figures 8.8a and 8.8b indicated that VSI (black line in Figure 8.8) is amongst the least contributing variables in generating a high Sharpe ratio. As seen in Figures 8.13a and 8.13b, the distribution of models containing VSI is not skewed in the validation set, however, seemingly seemingly positively skewed in the test set. The most prominent reasons for why the indicator is inefficient is assumed to be the problems described in Section 4.3; there are cases of corrections where VSI does not produce a signal, and also cases where the VSI produces signals outside corrections and crises. Furthermore, since it produces a signal only after volatility has increased in excess of 40 % in two days, a lag effect is introduced, impairing the correction response.



(a) Validation set



(b) Test set

Figure 8.13: Distribution of Sharpe ratios for models with VSI and without VSI in the validation and test set

Chapter 9

Conclusion

Our out-of-sample results show that selected financial variables carry predictive content for the occurrence of market downturns. With a realized Sharpe ratio of 0.86, our Candidate strategy outperforms the market at a significance level of 95 %. Thus, we find evidence against semi-strong form of market efficiency in the Norwegian stock market during the period from march 2014 to march 2019.

The results show that strategies encompassing the technical momentum-based CCI indicator notably outperforms the strategies that do not. The Price-to-book ratio less a 1-year rolling average show similar effects, however not as distinct as in the case of the CCI. Amongst the poorest performing variables, we find SX250 and the VSI. In terms of indicators, possible advancements include the addition of derivative markets data, as discussed in Li et al. (2015). The performance of the CCI variable used in this study suggests that other technical indicators may also be of interest, with examples including trading volume, support and resistance, and the Relative Strength Index (RSI).

The definitions presented in Chapter 3 represent a preliminary approach to generically defining the different states of the Norwegian equity market. Academic literature in this area may thus benefit from both refining the proposed definitions and developing alternative approaches. One possibility is to define overpricing more symmetrically around the price peak using a Hodrick-Prescott filter. Running a given classification model using several different definitions would also be of value in terms of model robustness testing.

There is a notable discrepancy in the model's ability to recognize the different states. Out-of-sample assessment shows that the model struggles to recognize impending market corrections, while promptly responding to ongoing price declines. Our study shows that the probabilities significantly contributes to a high Sharpe ratio. Further studies may benefit of exploring alternative classification models, where recent literature presents a variety of alternative approaches. One example is the use of neural network approaches, which Fioramanti (2008) finds outperforms more traditional methods under certain conditions. While the authors argue that the non-parametric nature of the approach impairs its value to policy makers, the results are compelling from a trading perspective.

Introducing a time-varying shrinkage factor represented by the estimated probability of being in a state of relative financial stability, we conclude that the Candidate strategy outperforms a portfolio solely based on the Kelly criterion when no gearing is allowed. This highlights some of the limitations of the Kelly criterion when applied to capital markets, as violations of notable assumptions causes the Kelly criterion to suggest exposures exceeding the risk preferences of a log utility investor. Further research within this field could benefit from utilizing alternative risk measures or return distributions, in the pursuit to create models that more closely resemble the characteristics of real financial markets.

Bibliography

Allen, F., Haas, M., Nowak, E., and Tengulov, A. (2018). Market efficiency and limits to arbitrage: Evidence from the volkswagen short squeeze.

Atkinson, P. M. and Massari, R. (2011). Autologistic modelling of susceptibility to landsliding in the central apennines, italy. *Geomorphology*, 130(1-2):55–64.

Baker, R. D. and McHale, I. G. (2013). Optimal betting under parameter uncertainty: improving the kelly criterion. *Decision Analysis*, 10(3):189–199.

Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of accounting research*, pages 159–178.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1):3–18.

Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2015). X-capm: An extrapolative capital asset pricing model. *Journal of financial economics*, 115(1):1–24.

Berg, A., Borensztein, E., and Pattillo, C. (2005). Assessing early warning systems: how have they worked in practice? *IMF staff papers*, 52(3):462–502.

Bernoulli, D. (1738). *Specimen theoriae novae de mensura sortis*. Gregg.

Bhojraj, S. and Lee, C. M. (2002). Who is my peer? a valuation-based approach to the selection of comparable firms. *Journal of accounting research*, 40(2):407–439.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association*, page 177–181.

Borch, B. N. (2008). Market efficiency theory and the earnings announcement premium at the oslo stock exchange. Master's thesis.

Breiman, L. et al. (1961). Optimal gambling systems for favorable games.

Brunnermeier, M. K. and Oehmke, M. (2012). Bubbles, finiancial crises, and systemic risk.

Bussiere, M. and Fratzscher, M. (2006). Towards a new early warning system of financial crises. *journal of International Money and Finance*, 25(6):953–973.

Casu, B., Clare, A., and Saleh, N. (2011). Towards a new model for early warning signals for systemic financial fragility and near crises: an application to oecd countries.

Christoffersen, P. (2011). *Elements of financial risk management*. Academic Press.

Chung, S.-L., Tsai, W.-C., Wang, Y.-H., and Weng, P.-S. (2011). The information content of the s&p 500 index and vix options on the dynamics of the s&p 500 index. *Journal of Futures Markets*, 31(12):1170–1201.

Coudert, V. and Gex, M. (2008). Does risk aversion drive financial crises? testing the predictive power of empirical indicators. *Journal of Empirical Finance*, 15(2):167–184.

Dalen, P. M. (2014). Er oslo børs et effisient aksjemarked?: en test av markedseffisiens på svak form i perioden 1996-2013. Master's thesis, Norwegian University of Life Sciences, Ås.

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

Daniel, K. D., Hirshleifer, D., and Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance*, 56(3):921–965.

Dimson, E. and Mussavian, M. (1998). A brief history of market efficiency. *European financial management*, 4(1):91–103.

Douady, R. and Kornprobst, A. (2018). An empirical approach to financial crisis indicators based on random matrices. *International Journal of Theoretical and Applied Finance*, 21(03):1850022.

Drees, M. B. (2005). *Asset mispricing due to cognitive dissonance.* International Monetary Fund.

Engebretsen, J. (2017). Analyse av unormal kursendring ved annonsering av oppkjøp på oslo børs. Master's thesis, Norwegian University of Life Sciences, Ås.

Engevik, Ø. and Helleren, O. A. (2009). Innsidehandel: en empirisk studie av oslo børs. Master's thesis.

Fioramanti, M. (2008). Predicting sovereign debt crises using artificial neural networks: a comparative approach. *Journal of Financial Stability*, 4(2):149–164.

Flahaut, B. (2004). Impact of infrastructure and local environment on road unsafety: Logistic modeling with spatial autocorrelation. *Accident Analysis & Prevention*, 36(6):1055–1066.

Frankel, J. and Saravelos, G. (2012). Can leading indicators assess country vulnerability? evidence from the 2008–09 global financial crisis. *Journal of International Economics*, 87(2):216–231.

Frankel, J. A. and Rose, A. K. (1996). Currency crashes in emerging markets: An empirical treatment. *Journal of international Economics*, 41(3-4):351–366.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of financial economics*, 8(1):55–69.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics New York.

Fu, J., Zhou, Q., Liu, Y., and Wu, X. (2019). Predicting stock market crises using daily stock market valuation and investor sentiment indicators. *The North American Journal of Economics and Finance.*

Fuertes, A.-M. and Kalotychou, E. (2006). Early warning systems for sovereign debt crises: The role of heterogeneity. *Computational statistics & data analysis*, 51(2):1420–1441.

Gujarati, D. and Porter, D. (2003). Multicollinearity: What happens if the regressors are correlated. *Basic econometrics*, 363.

Hagestande, E. and Hals, M. (2012). Emisjonsannonseringer på oslo børs: en event-studie. Master's thesis.

Hamre, M. F. and Sande, A. (2014). Kan man oppnå meravkastning ved å følge innsidere på oslo børs?: en kvantitativ analyse av finansavisens innsideportefølje. Master's thesis.

Harmon-Jones, E. and Mills, J. (1999). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory.

Haug, M. and Hirschey, M. (2006). The january effect. *Financial Analysts Journal*, 62(5):78–88.

Haugen, R. A. and Jorion, P. (1996). The january effect: Still there after all these years. *Financial Analysts Journal*, 52(1):27–31.

Hermanrud, P. (2019). Meeting with sparebank 1 markets.

Herwartz, H. and Kholodilin, K. A. (2014). In-sample and out-of-sample prediction of stock market bubbles: Cross-sectional evidence. *Journal of Forecasting*, 33(1):15–31.

Herwartz, H. and Morales-Arias, L. (2009). In-sample and out-of-sample properties of international stock return dynamics conditional on equilibrium pricing factors. *The European Journal of Finance*, 15(1):1–28.

Holmes, M. J. and Otero, J. (2017). Re-examining the movements of crude oil spot and futures prices over time. *Energy Economics*.

Hovland, K. M. (2017). Statoil kutter mer enn ventet.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Jegadeesh, N. and Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of finance*, 56(2):699–720.

Jegadeesh, N. and Titman, S. (2011). Momentum. *Annu. Rev. Financ. Econ.*, 3(1):493–509.

Kadane, J. B. (2011). Partial-kelly strategies and expected utility: Small-edge asymptotics. *Decision Analysis*, 8(1):4–9.

Kaizoji, T. and Sornette, D. (2008). Market bubbles and crashes.

Kaminsky, G., Lizondo, S., and Reinhart, C. M. (1998). Leading indicators of currency crises. *Staff Papers*, 45(1):1–48.

Kara, Y., Boyacioglu, M. A., and Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert systems with Applications*, 38(5):5311–5319.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of financial economics*, 12(1):13–32.

Kelly Jr, J. L. (2011). A new interpretation of information rate. In *The Kelly Capital Growth Investment Criterion: Theory and Practice*, pages 25–34. World Scientific.

Kim, K.-j. and Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, 19(2):125–132.

Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., and Klein, M. (2002). *Logistic regression*. Springer.

Kordos, M. and Cwiok, A. (2011). A new approach to neural network based stock trading strategy. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 429–436. Springer.

Kornprobst, A. (2017). Winning investment strategies based on financial crisis indicators.

Lamache, A. T. and Sandøy, E. V. (2016). Statistisk arbitrasje: kan pairs trading avkrefte svak markedseffektivitet på oslo børs? Master's thesis.

Lambert, D. R. (1983). Commodity channel index: Tool for trading cyclic trends. *Technical Analysis of Stocks & Commodities*, 1:47.

Langli, J. V. (2015). Innsidehandel og asymmetrisk informasjon på oslo børs: en begivenhetsstudie av markedsreaksjoner på meldepliktig innsidehandel. Master's thesis.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Li, W.-X., Chen, C. C.-S., and French, J. J. (2015). Toward an early warning system of financial crises: What can index futures and options tell us? *The Quarterly Review of Economics and Finance*, 55:87–99.

Lie, E. and Lie, H. J. (2002). Multiples used to estimate corporate value. *Financial Analysts Journal*, 58(2):44–54.

Liu, J., Nissim, D., and Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research*, 40(1):135–172.

Long, J. S. (1997). Regression models for categorical and limited dependent variables (vol. 7). *Advanced quantitative techniques in the social sciences.*

Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. *The Journal of finance*, 50(2):549–572.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Mamelund, K. (2006). Kortsiktig overreaksjon ved store kursendringer på oslo børs. Master's thesis.

Manasse, P., Roubini, N., and Schimmelpfennig, A. (2003). Predicting sovereign debt crises, imf working paper no. wp/03/221, international monetary fund, washington, dc.

Markovitz, H. (1959). Portfolio selection: Efficient diversification of investments. *NY: John Wiley.*

Markwat, T., Kole, E., and Van Dijk, D. (2009). Contagion as a domino effect in global stock markets. *Journal of Banking & Finance*, 33(11):1996–2012.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257.

Mishkin, F. S. and White, E. N. (2002). Us stock market crashes and their aftermath: implications for monetary policy. Technical report, National bureau of economic research.

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2009). What factors affect the oslo stock exchange?

Nekrasov, V. (2014). Kelly criterion for multivariate portfolios: A model-free approach. *Available at SSRN 2259133*.

Nyhus, V. and Skjørten, E. B. (2016). En test av optimerte handelsstrategier på oslo børs. Master's thesis.

Olafsson, S. E. and Fossen, S. (2014). Oppkjøpsannonseringer på oslo børs: en event-studie med fokus på kjøpers ekstraordinære avkastning. Master's thesis.

Ørpetveit, A. and Hansen, T. (2016). Markedseffisiens på oslo børs: en studie av prisanomalier før, under og etter finanskrisen. Master's thesis.

Patel, J., Shah, S., Thakkar, P., and Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, 42(4):2162–2172.

Patel, S. A. and Sarkar, A. (1998). Crises in developed and emerging stock markets. *Financial Analysts Journal*, 54(6):50–61.

Peter, M. (2002). Estimating default probabilities of emerging market sovereigns: A new look at a not-so-new literature. *Available at SSRN 1017468*.

Peterson, Z. (2017). Kelly's criterion in portfolio optimization: A decoupled problem.

Philpot, J. and Peterson, C. A. (2011). A brief history and recent developments in day-of-the-week effect literature. *Managerial Finance*, 37(9):808–816.

Rabin, M. (1998). Psychology and economics. *Journal of economic literature*, 36(1):11–46.

Rozeff, M. S. and Kinney Jr, W. R. (1976). Capital market seasonality: The case of stock returns. *Journal of financial economics*, 3(4):379–402.

Sand, M. S. and Bødal, F. W. (2013). Markedseffisiens og asymmetrisk informasjon på oslo børs: En begivenhetsstudie av markedets reaksjoner på resultatvarsel i perioden 01.01. 2000 til 31.12. 2012. Master's thesis.

Savona, R. and Vezzoli, M. (2015). Fitting and forecasting sovereign defaults using multiple risk signals. *Oxford Bulletin of Economics and Statistics*, 77(1):66–92.

Schreiber-Gregory, D. (2018). Logistic and linear regression assumptions: Violation recognition and control.

Schwab Center for Financial Research (2018). Market correction: What does it mean? Retrieved 2019-05-29.

Schwert, G. W. (2003). Anomalies and market efficiency. *Handbook of the Economics of Finance*, 1:939–974.

Sewell, M. (2007). Behavioural finance. *University of Cambridge*, pages 1–14.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1):119–138.

Simonsen, L.-C. (2012). Er det norske aksjemarkedet effisient?: en test av svak form for markedseffisiens gjennom bruk av en inter-dag strategi. Master's thesis, Universitetet i Tromsø.

Theodossiou, P. and Lee, U. (1993). Mean and volatility spillovers across major national stock markets: Further empirical evidence. *Journal of Financial Research*, 16(4):337–350.

Thorp, E. O. (2011). The kelly criterion in blackjack sports betting, and the stock market. In *The Kelly Capital Growth Investment Criterion: Theory and Practice*, pages 789–832. World Scientific.

Tollefsen, T. (2010). Er oslo børs svakteffisient?: en test av teknisk analyse. Master's thesis.

Van Dijk, M. A. (2011). Is size dead? a review of the size effect in equity returns. *Journal of Banking & Finance*, 35(12):3263–3274.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

Wu, W.-T. A. (2014). The p/e ratio and profitability. *Journal of Business & Economics Research (Online)*, 12(1):67.

Yu, L., Wang, S., and Lai, K. K. (2005). Mining stock market tendency using ga-based support vector machines. In *International Workshop on Internet and Network Economics*, pages 336–345. Springer.

Ziemba, W. T. et al. (2003). *The stochastic programming approach to asset, liability, and wealth management*. Research Foundation of AIMR, Scorpion Publications.

Zulovic, K. (2012). Annonseringer av innsiderhandler på oslo børs - en test av semi-sterk form markedseffisiens. Master's thesis, Norges teknisk-naturvitenskapelige universitet.

Appendices

Appendix A

Crisis and Corrections

Occurrences of Crises and Corrections

Start Date	End Date	Duration	% Decline
Corrections			
1984-05-09	1984-07-02	36	20.9 %
1985-02-08	1985-03-26	33	10.2 %
1987-04-24	1987-05-06	8	10.0 %
1987-09-21	1987-11-02	30	29.7 %
1989-09-28	1989-10-16	13	14.2 %
1990-03-16	1990-04-26	27	9.1 %
1990-08-02	1990-09-13	30	9.3 %
1993-08-13	1993-09-09	20	8.4 %
1997-02-20	1997-04-03	28	8.5 %
1997-10-22	1997-12-22	44	13.2 %
1998-05-06	1998-06-19	30	11.5 %
1999-09-20	1999-10-18	21	9.2 %
2000-03-07	2000-04-17	29	16.2 %
2000-09-14	2000-10-26	30	7.2 %
2003-09-04	2003-09-30	19	8.9 %
2004-03-08	2004-05-10	43	12.0 %
2005-04-08	2005-04-28	15	8.1 %
2005-10-03	2005-10-19	13	15.5 %
2006-05-11	2006-06-13	21	20.3 %
2007-02-14	2007-03-14	21	8.2 %
2007-07-19	2007-08-16	21	16.0 %
2010-01-11	2010-02-15	26	11.4 %
2010-04-26	2010-07-01	46	19.2 %
2018-09-25	2018-12-27	65	17.2 %
Crises			
1987-09-21	1987-12-16	63	45.4 %
1990-08-02	1992-08-25	512	55.0 %
1998-05-06	1998-10-08	110	46.2 %
2000-09-14	2003-02-26	613	59.1 %
2008-05-22	2008-11-21	132	64.0 %

Table A.1: Occurrences of crises and corrections on the OSE in the time period 1983-2019

Appendix B

Variables

Price-to-Book Ratio (PB)

The PB ratio equals the total market capitalization of OSEBX divided by the combined book value of equity of all companies constituting the index. As the book value of equity represents the shareholders' claim in the company after liabilities have been covered, the PB reflects the relative price of a company's equity. Daniel et al. (2001) states that the indicator catches a combination of market risk and mispricing. Consequently, the hypothesis is that if the price of the index drifts considerably from its underlying equity value, this would suggest increased risk of stock market overvaluation. The explicit formula is stated in equation B.1.

$$\mathrm{PB}_t = \frac{\mathrm{OSEBX}_t}{\mathrm{Book}_t} \tag{B.1}$$

Several studies argue that the price-to-book ratio carry information content for the incident of overpricing. In a study of predicting stock market bubbles, Herwartz and Kholodilin (2014) found PB to be the best in- and out-of-sample indicator of an impending asset price bubble, while Fu et al. (2019) also conclude its position as a suitable indicator of stock market bubbles. Earlier studies such as Lie and Lie (2002), who assessed the relative performance of various multiples in asset valuation, found that the PB multiple usually generates less biased estimates of asset values than sales- and earnings multiples.

Forward Price-Earnings Ratio (PE)

Another commonly used financial valuation multiple is the price-earnings (PE) ratio. The ratio relates stock value to the actual profits of a company, thus reflecting the market's expectation of firm risk and future growth (Wu, 2014). The hypothesis is that a high PE ratio might indicate overconfidence in growth estimates and, consequently, increased risk of an asset price bubble. Adversely, a low PE ratio could signal underconfidence and, subsequently, a buying opportunity. The OSEBX PE used in this thesis is the weighted average PE of the companies listed on the stock exchange, and is found by dividing the price by the Earnings Per Share (EPS).

$$\mathrm{PE}_t = \frac{\mathrm{OSEBX}_t}{\mathrm{EPS}_t} \tag{B.2}$$

The value of a company's equity, and thus its stock price, should theoretically equal all future cash flows discounted at an appropriate discount rate. The stock price itself should therefore reflect future, rather than current, earnings. A disparity between the stock price and estimated future earnings may therefore better reflect overpricing relative to fundamental value compared to current earnings. Such a rationale is consistent with the findings of both Liu et al. (2002) and Lie and Lie (2002), whom both concluded that multiples derived from forward earnings outperform multiples derived from historical earnings.

PE have been used to great effect in monitoring the possible occurence of a price bubble. Both Coudert and Gex (2008) and Fu et al. (2019) found that an increasing PE ratio lead to a statistically significant increase in the probability of stock market crisis.

EQNR / 3Y Brent future (EQ)

'The Oil price significantly affects cash flows of most industry sectors at OSE' (Næs et al., 2009). A ratio of OSEBX relative to the oil price may therefore indicate over- and overvaluation. However, the OSEBX is comprised of companies from a wide range of sectors, and the ratio may therefore not be immunized against drifts in the oil price. The pricing of OSEBX relative to oil may therefore be a noisy indicator.

The goal is to identify factors which are equally impacted by changes in fundamentals, and consequently, represents more or less stable equilibrium ratios. As evident from Figure B.1, Equinor ASA (EQNR) has constituted a large share of the total market cap of OSEBX. Furthermore, EQNR, being an energy company heavily focused on oil, is expected to exhibit analogous reactions to fundamental trends affecting drift in oil price. Consequently, the relative pricing of EQNR and Crude oil is expected to be largely immunized against stochastic trends and subsequently be a better indicator of impending asset price bubbles on OSEBX.

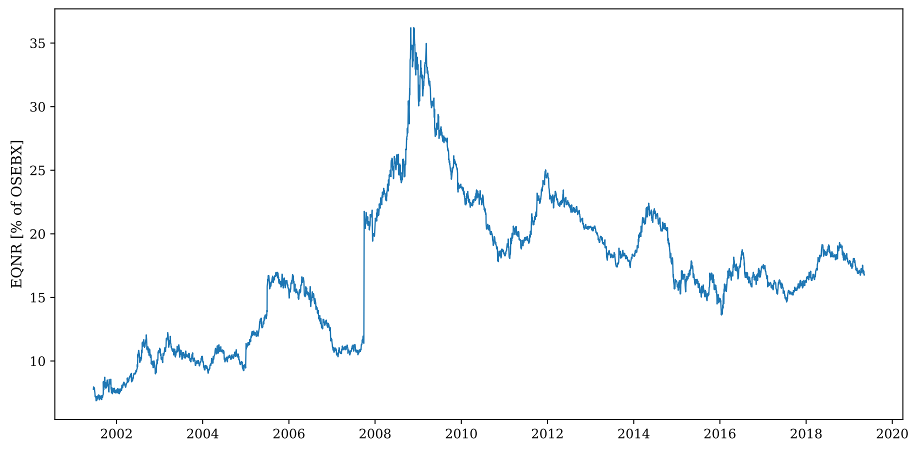


Figure B.1: Weighting of EQNR in the OSEBX index

Both spot and futures are possible representatives for the the oil price. The Samuelson effect states that contracts close to maturity react more to given information compared to contracts with long time to maturity (Holmes and Otero, 2017). Long-term contracts are therefore likely to be less affected by short-term noise, and consequently assumed to be more representative for EQNR's value. Due to data quality, long-term contracts such as 4 and 5 years are not suitable for our study. They have historically been trading more infrequent compared to shorter, leading to long periods of constant values. This also applies to the 3Y contract, however only up to 27.02.2006. The 1 month contracts (1M) have been traded daily at least since 1998. The 3Y and 1M are therefore linked 27.02.2006, which means that values before this date are 1M, and 3Y thereafter. Given the 6 % difference between 1M and 3Y on the linkage day, some boundary effects will occur. Figure B.2 shows the linked contract, and Formula B.3 shows the indicator used.

Figure B.2: Brent 3Y future and Linked 3Y1M Brent future

$$\mathrm{EQ}_t = \frac{\mathrm{EQNR}_t}{\mathrm{3YBrent}_t} \tag{B.3}$$

Price-Earnings OSEBX/STOXX (SX)

This variable reflects the relationship between the PE ratios of the OSEBX and the STOXX Europe 600 index (formula B.4). By comparing the PE, potential distinctive pricing may be identified and in turn signal over- or undervaluation.

$$\mathrm{SX}_t = \frac{\mathrm{PEOSEBX}_t}{\mathrm{PESTOXX}_t} \tag{B.4}$$

A potential drawback of this indicator originates in different sectoral composition of OSEBX and STOXX 600. Fundamental drivers may accordingly affect the numerator and denominator of the indicator disproportionally. As a result, the indicator may not be immunized against stochastic trends, and thus be less efficient in signaling an impending market corrections. This is supported by Bhojraj and Lee (2002), which state that the accuracy of PE is better when comparing to industry peers. The next variable mitigates this concern by having the same industry composition as OBX.

OBX / Synthetic Index (SY)

The synthetic OBX index developed by Sparebank 1 Markets is a hypothetical portfolio of foreign stocks aimed at replicating the OBX index with the same industry composition. One drawback, however, is the potential difficulties in obtaining suitable replacements for industries specific to Norway, the most prominent example being salmon farming (Hermanrud, 2019). Figure B.3 shows the value of OBX relative OSEBX, and substantiate the assumption of OBX being largely representative for OSEBX (average of 88 % for the test period). Formula B.5 shows the equation for SY.

$$\mathrm{SY}_t = \frac{\mathrm{OSEBX}_t}{\mathrm{Synthetic}_t} \tag{B.5}$$

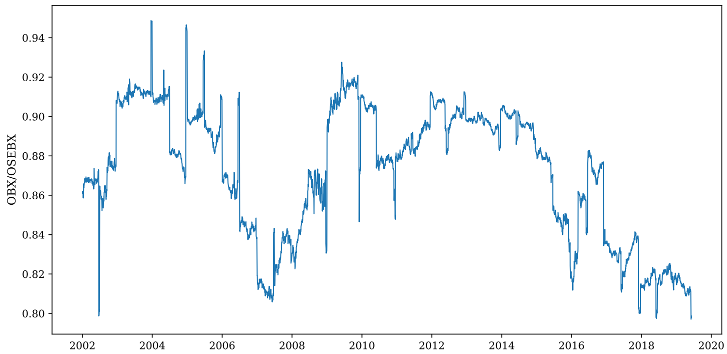


Figure B.3: Value of OBX as share of OSEBX

CBOE Volatility Index (VIX)

The CBOE Volatility Index (VIX) is the volatility implied by S&P 500 index options, and is defined as the expected short-term (30 days) volatility. Several previous studies, such as Coudert and Gex (2008), Chung et al. (2011) and Li et al. (2015), conclude that the VIX and implied volatility provide explanatory power for the risk of crises.

Commodity Channel Index (CCI)

CCI is a momentum indicator aiming at recognizing significant price deviations from a recent mean. Using a 120 days lag, CCI is calculated by the following formula (Lambert, 1983):

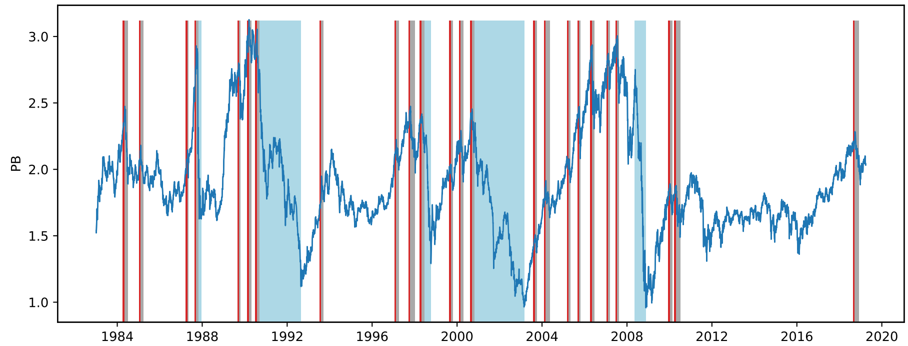
$$\begin{aligned} \mathrm{CCI}_t &= \frac{\mathrm{OSEBX}_t - \mathrm{SMA}_{t-j}}{MAD_{t-j}} \\ \mathrm{CCI}_t &= \frac{\mathrm{OSEBX}_t - \frac{1}{\tau+1}\sum_{j=0}^{\tau}\mathrm{OSEBX}_{t-j}}{\frac{1}{\tau+1}\sum_{j=0}^{\tau}|\mathrm{OSEBX}_t - \frac{1}{\tau+1}\sum_{0=1}^{\tau}\mathrm{OSEBX}_{t-j}|} \end{aligned} \tag{B.6}$$

where SMA refers to Simple Moving Average and MAD is the Mean Absolute Deviation.

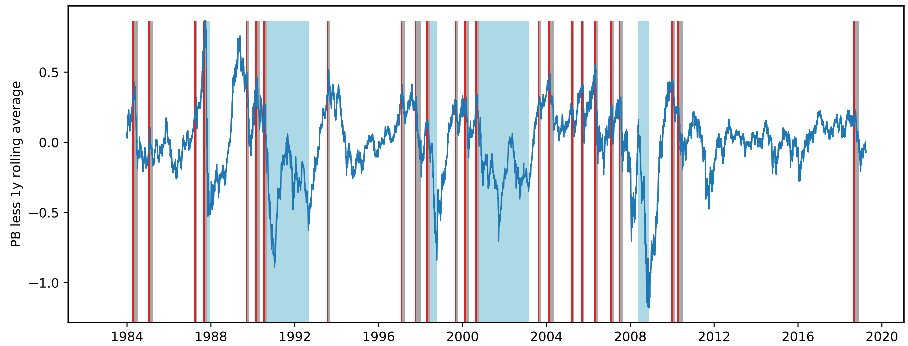
List of Variables with Sources

Variable (Source)	Description	Available from
OSEBX (Oslo Stock Exchange)	Price of Oslo Stock Exchange Benchmark Index	1983
Book (Sparebank 1 Markets)	Book value of OSEBX	1983
EPS OSEBX (Sparebank 1 Markets)	12 month Earnings-per-share future of OSEBX	2001
EPS STOXX (Sparebank 1 Markets)	12 month Earnings-per-share future of STOXX	2000
EQNR (Sparebank 1 Markets)	Equinor stock price adjusted for stock splits and dividends	2001
3Y Brent (Macrobond)	3-year future contract on Brent crude oil	1998
Synthetic OBX index (Sparebank 1 Markets)	Synthetic industry-neutral index mimicing OSEBX	1997
CBOE VIX (Fred St. Louis)	Implied volatility of S&P 500	1990

Plots of Original and Transformed Variables

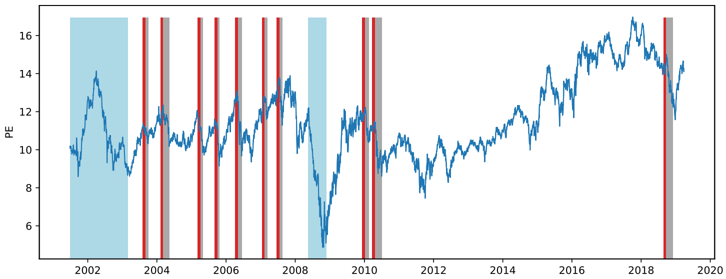


(a) Price-to-book ratio of OSEBX

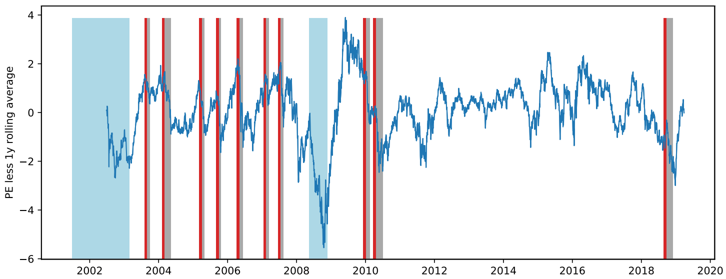


(b) Price-to-book ratio less 1y rolling average

Figure B.4: Original and transformed Price-to-book ratio of OSEBX

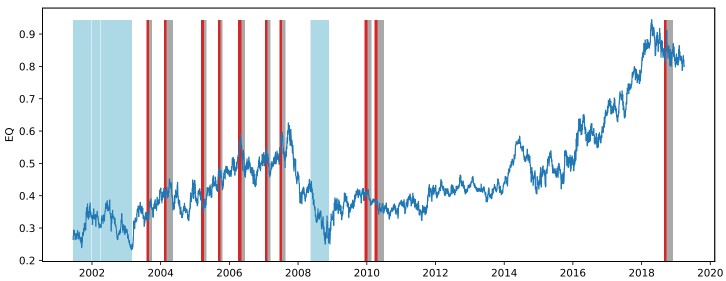


(a) Price-earnings ratio of OSEBX

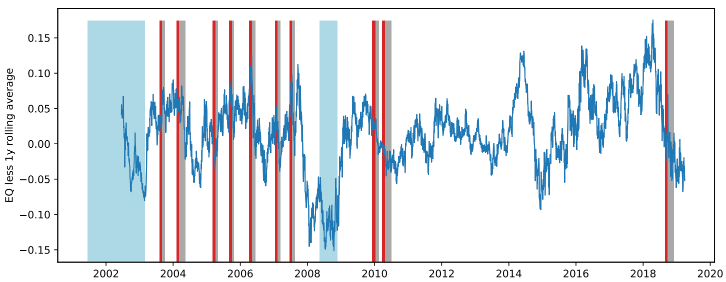


(b) Price-earnings ratio less 1y rolling average

Figure B.5: Original and transformed Price-earnings ratio of OSEBX

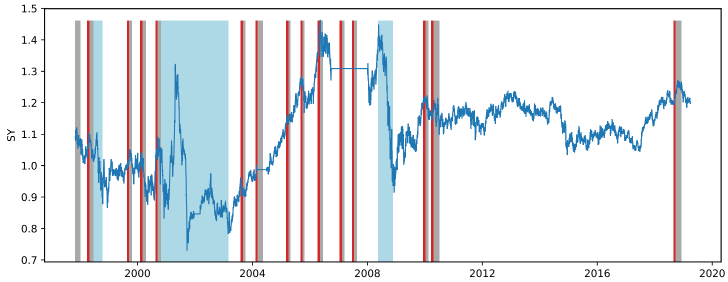


(a) EQNR / 3y Brent

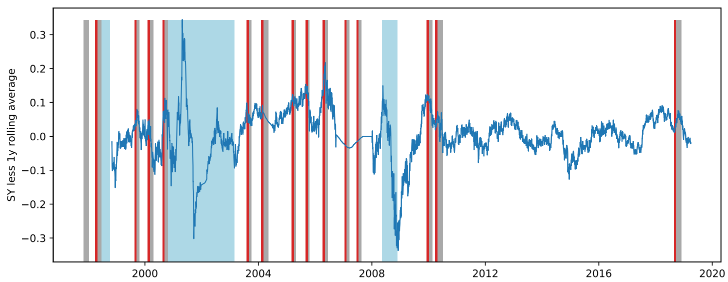


(b) EQNR / 3y Brent less 1y rolling average

Figure B.6: Original and transformed EQNR / 3y Brent of OSEBX

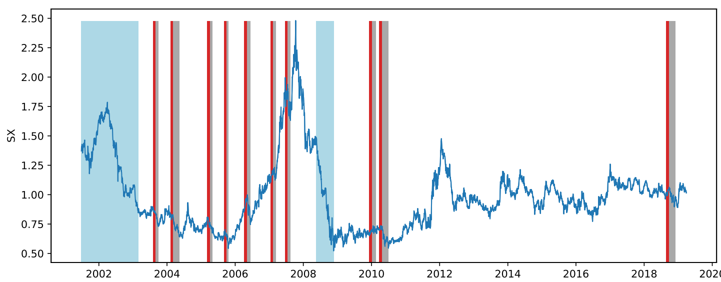


(a) OSEBX/Synthetic

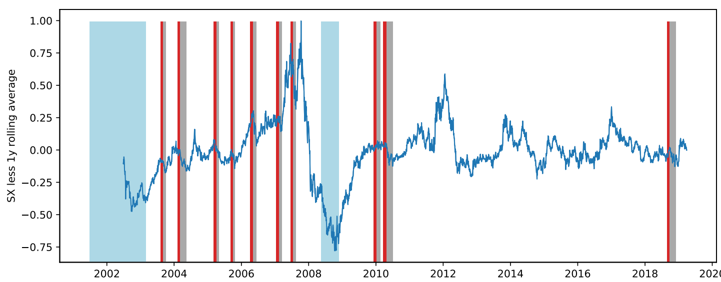


(b) OSEBX/Synthetic less 1y rolling average

Figure B.7: Original and transformed OSEBX / Synthetic index



(a) PE OSEBX / PE STOXX



(b) PE OSEBX / PE STOXX less 1y rolling average

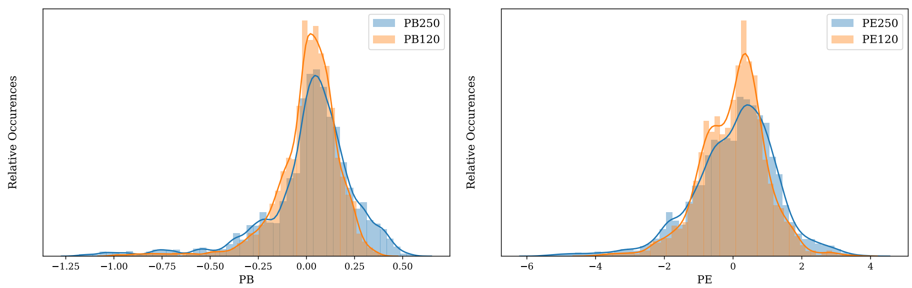
Figure B.8: Original and transformed PE OSEBX / PE STOXX

Descriptive Statistics

The descriptive statitics of these variables are included in the table below.

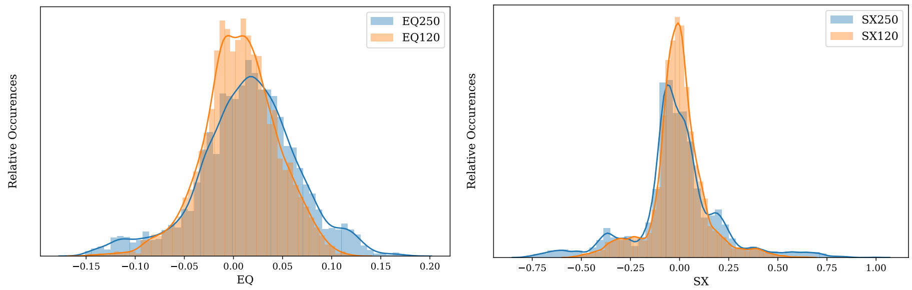
	Mean	Standard deviation	Min	Max	Skewness	Excess kurtosis	Num. observations
PB250	0.0019	0.258	-1.18	0.87	-0.67	1.79	8848
PB120	0.0006	0.184	-1.00	0.64	-1.04	3.20	8978
PE250	0.0482	1.235	-5.55	3.90	-0.67	1.59	4211
PE120	0.0456	0.972	-4.12	3.66	-0.30	0.96	4341
EQ250	0.0150	0.051	-0.15	0.18	-0.32	0.63	4220
EQ120	0.0071	0.037	-0.15	0.12	-0.22	0.46	4350
SX250	-0.0164	0.219	-0.78	0.99	0.11	2.46	4211
SX120	-0.0049	0.143	-0.55	0.64	0.03	2.29	4341
SY250	0.0050	0.069	-0.34	0.34	-0.64	3.66	5124
SY120	0.0019	0.057	-0.37	0.34	-0.98	8.25	5254
CCI	0.7075	1.683	-6.86	4.86	-0.66	0.03	8978
VIX	19.251	7.816	9.14	80.86	2.10	7.71	7340
VSI	0.0392	0.069	0.001	0.66	2.69	10.26	9092

Table B.1: Descriptive statistics



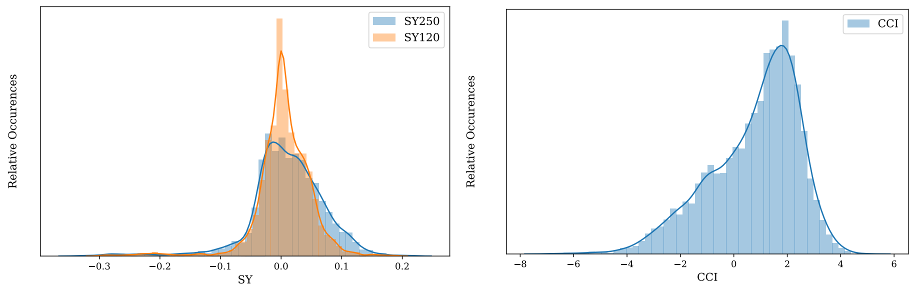
(a) PB (b) PE

Figure B.9: Distribution of transformed PB and PE



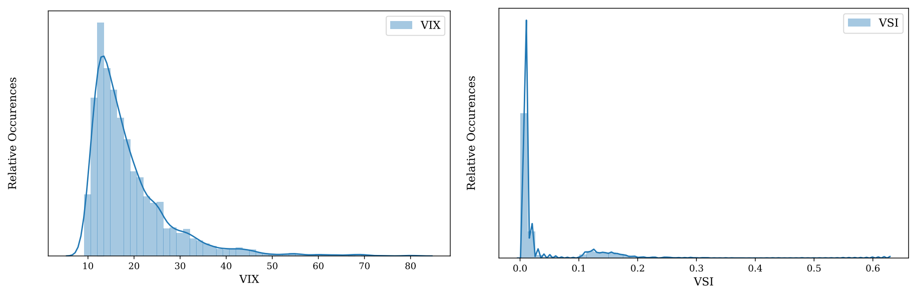
(a) EQ (b) SX

Figure B.10: Distribution of transformed EQ and SX



(a) SY (b) CCI

Figure B.11: Distribution of transformed SY and CCI



(a) VIX (b) VSI

Figure B.12: Distribution of VIX and VSI

Appendix C

Model Selection

Derivation of Maximum log likelihood

Consider a set of N mutually exclusive binary dependent variables each following a Bernoulli distribution, taking the general form

$$I_{it} \sim \text{Bernoulli}(\pi_{it}) \tag{C.1}$$

where I_{it} is assumed to be independent of observation t. Letting $i \in \{1, ..., \text{N}\}$ represent the possible states, and using that $\sum_{i=1}^{N} I_{it} = 1$, the probability distribution for observation t is

$$f(I_{it}, \boldsymbol{\beta}) = \prod_{i=1}^{\text{N}} \pi_{it}^{I_{it}} \tag{C.2}$$

The conditional probability function for T independent observations is given by

$$l(I_{it}, \boldsymbol{\beta}) = \prod_{t=0}^{\text{T}} \prod_{i=1}^{\text{N}} \pi_{it}^{I_{it}} \tag{C.3}$$

Taking the log for numerical considerations, we obtain the log likelihood function

$$\mathscr{L}(I_{it}, \boldsymbol{\beta}) = \sum_{t=0}^{\text{T}} \ln \left[\prod_{i=1}^{\text{N}} \pi_{it}^{I_{it}} \right] \tag{C.4}$$

Stationarity Test Results

Variables	Test statistic	p-value
Original		
PB	-1.94	0.32
PE	-2.61	0.09
EQ	-1.18	0.68
SX	-2.51	0.11
SY	-2.01	0.28
CCI	-6.40	2.1×10^{-8}
VSI	-8.05	1.8×10^{-12}
VIX	-4.21	6.0×10^{-4}
Transformed		
PB250	-3.3556	0.0126
PB120	-4.87	$4.1\ \times10^{-5}$
PE250	-4.3623	0.0003
PE120	-5.85	$3.6\ \times10^{-7}$
EQ250	-4.7770	6.02×10^{-5}
EQ120	-6.59	6.8×10^{-9}
SX250	-3.3115	0.0255
SX120	-4.81	5.2×10^{-5}
SY250	-3.6912	0.0043
SY120	-5.82	4.2×10^{-7}

Table C.1: Augmented Dickey-Fuller test for stationarity

Critical values at 1%, 5% and 10% confidence are -3.43, -2.86 and -2.57 respectively. Based on the test statistics and p-values, we reject the null hypothesis of CCI, VSI and VIX containing unit roots and consequently conclude that they are stationary processes. The null hypothesis is rejected for all transformed variables and we conclude that these are all stationary.

Appendix D

Trading Strategies

Estimation of Kelly Parameters

The expected $\hat{\mu}$ drift and volatility $\hat{\sigma}^2$ of the risky asset must be estimated. Equation 6.2 can be used to relate the drift and volatility of the risky asset to historical market data. Letting Avg[R] and Var[R] denote the estimated return and variance from the sample data, it can be shown (see e.g. Peterson (2017)) that:

$$\hat{\mu} = \ln\left(1 + \text{avg}[R]\right) \tag{D.1}$$

$$\hat{\sigma} = \ln\left(\text{var}[R]\exp(-2\hat{\mu}) + 1\right)^{\frac{1}{2}} \tag{D.2}$$

The average return is calculated by finding the average return from 04.01.1983 to the end of the training set (08.02.2010). The reason for using data back to 1983, is that a longer dataset is assumed to be more representative of the true drift, since a short period is more sensitive to periodical effects.

EGARCH (1,1) is used as estimator for the volatility. Christoffersen (2011) argues that EGARCH have the advantages of capturing the leverage effect described in Black (1976) and a logarithmic specification ensuring the variance is always positive. A time-varying volatility is used, because it is assumed to be a better approximation of the one-step volatility. By estimating the coefficients, we obtain the following volatility estimator:

$$\ln(\mathrm{var}[R]) = \omega + \alpha(\phi R_t + \gamma[|R_t| - E|R_t|]) + \beta \ln(\mathrm{var}[R])^2 \tag{D.3}$$

Parameter	**Value**
ω	-0.412
α	-0.0683
ϕ	0.000904
γ	0.295
β	0.953

Table D.1: Estimated parameters of EGARCH (1,1) volatility model

NTNU
Norwegian University of
Science and Technology