Norwegian School of Economics
Bergen, Fall 2020

Economic Policy Uncertainty and Norwegian Stock Returns

An empirical study of the relation between economic policy uncertainty and stock returns with evidence from the Oslo Stock Exchange in the period of 1992 - 2019

Sindre Kolnes Eriksen & Jens Sønsteby Tobiassen

Supervisor: Nataliya Gerasimova

Master Thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This study investigates if economic policy uncertainty (EPU) is priced in stock returns at the Oslo Stock Exchange (OSE). The analysis is conducted by exploring the linear relationship between exposure to economic policy uncertainty and expected stock returns through the Fama-Macbeth framework. The estimates obtained are controlled for several well-renowned factor pricing models to isolate the policy uncertainty effect. We apply four different methods of capturing economic policy uncertainty to increase the robustness of the analysis. This includes measures based on (i) print newspaper articles, (ii) online newspaper articles, (iii) Google searches and (iv) a firm-specific measure obtained by applying textual analysis to annual reports.

We do not find evidence of a negative linear relationship between economic policy uncertainty and expected stock returns. The extreme portfolios sorted by EPU exposure do not obtain significantly different return spreads. When controlling for the CAPM and the multi-factor models in context of the Fama-Macbeth framework, our portfolios obtain insignificant risk premia estimates associated with economic policy uncertainty. However, we do obtain significant estimates at one sorting method for two model specifications when applying the Google search-based measure of economic policy uncertainty. Nevertheless, the evidence is considered too limited for economic policy uncertainty to acquire status as a systematic risk factor in Norwegian stock returns.

Preface

This thesis is written as a part of our Master of Science in Economics and Business administration with specialization in Financial Economics at the Norwegian School of Economics (NHH).

Our interest for the subject was motivated by courses at NHH, Nova School of Business and Economics and University of Queensland, combined with our interest in financial markets and its interaction with macroeconomics. We hope our work may be of interest to academics as well as practitioners in the Norwegian stock market.

Writing this thesis has been a challenging exercise which has led to great learning, especially regarding working with large sets of data. We have developed an understanding and respect for academic research and the effort required to present it in a structured manner.

We would like to express our gratitude to our supervisor, Assistant Professor Nataliya Gerasimova, who has helped us along the way providing valuable guidance and feedback. Additionally, we would like to thank Thomas Katralen, Customer Manager at Atekst Retriever, for clarification and help when manoeuvring the database. Finally, we would like to thank our lecturers and fellow students for making our years at NHH highly educational, challenging and motivating.

Table of Contents

TABLE OF CONTENTS **4**

1 INTRODUCTION **5**

2 LITERATURE REVIEW **7**

2.1 AN INTRODUCTION TO POLICY UNCERTAINTY 7

2.2 ASSET PRICING 10

3 METHODOLOGY **15**

3.1 CREATING EPU INDICES 15

3.2 FACTOR PRICING 22

4 DATA **34**

4.1 EPU DATA 35

4.2 STOCK DATA 37

5 RESULTS **42**

5.1 OVERVIEW OF THE EPU FACTOR 43

5.2 FACTOR REALIZATIONS 46

5.3 FACTOR EXPOSURES 48

5.4 EPU AS FACTOR PREMIA 56

6 CONCLUSION **63**

REFERENCES **65**

APPENDIX **71**

R SCRIPT **90**

1 Introduction

Governments set the rules of the game. Political actions have substantial impacts on financial markets with many events of purely political nature leading to profound market reactions. Examples are Brexit in 2016 and the ongoing trade war between the United States and China which started in 2018 and caused turbulence in stock markets around the world. Not only do governments form policies that affect business conditions, but they are also one of the largest agents in economies, with expenditures constituting a sustainable share of gross domestic product. In literature, no doubt exists regarding the importance of government policy on the business environment (Friedman, 1968; Rodrik, 1991). However, the empirical implications have not been investigated until recently, much credited to Baker et al. (2013) who introduced a method of quantifying policy uncertainty as well as making it publicly available. This led to great attention in empirical research, with the findings that policy uncertainty has real implications on economic agents (Colak et al., 2017; Jens, 2017; Bonaime et al., 2018; Walkup, 2016). In this thesis, we use Al-Thaqeb and Algharabali's (2019) definition of policy uncertainty and define the variable as: "the economic risk associated with undefined future government policies and regulatory frameworks". Note that the terms economic policy uncertainty, EPU and policy uncertainty are used interchangeably throughout the thesis.

The literature is not limited to the investigation of changes in firm behaviour as a consequence of policy uncertainty. Brogaard and Detzel (2014) investigate the role of policy uncertainty in the cross-section of U.S. stock returns and find it commands a significant negative risk premium. This thesis examines the relation between exposure to economic policy uncertainty and expected stock returns at the OSE. We investigate this by testing the hypothesis: *economic policy uncertainty carries a negative risk premium in the Norwegian stock market*. We apply a dataset of stock returns, accounting data and four distinct measures of policy uncertainty to investigate if economic policy uncertainty is a systematic risk factor in the Norwegian stock market. It is particularly interesting to investigate the role of economic policy uncertainty in the cross-section of Norwegian stock returns as the Norwegian government constitutes an above average large part of the economy. While the average share of government expenditures in relation to mainland GDP was 44.6% in OECD countries in 2017, it was 58.1% in Norway (Riekeles, 2017). Consequently, governmental policy in Norway could be influencing economic agents more than in other countries.

We find little evidence of policy uncertainty carrying a significant risk premium in the cross-section of stock returns. This is the case both when comparing return spreads on extreme portfolios sorted on EPU and when controlling for the CAPM, the Fama-French Three-factor model and the Fama-French Five-factor model. We apply the Fama-Macbeth framework using three different sorting mechanisms when estimating risk premia. First, test portfolios' factor loadings are estimated through time series regressions before estimated risk premia are obtained by cross-sectional regressions. We obtain significant estimates at one sorting method for two model specifications when applying one of our measures of policy uncertainty: the Google search-based measure. However, the evidence is considered too limited in regard to economic policy uncertainty obtaining a foothold as a systematic risk factor in the Norwegian stock market.

Most papers concerned with the role of economic policy uncertainty in the cross-section of stock returns apply the method of Baker et al. (2013). However, as no universal way of capturing EPU is established in literature, we implement several measures different from the method of Baker et al. (2013) in order to increase the robustness of the analysis. In total, we apply four different methods to the analysis, where three are aggregate time series and the fourth is a firm-specific measure. The first is a measure following the methodology of Baker et al. (2013). We create an index by measuring the relative frequency of monthly print newspaper articles concerned with EPU. The second is a corresponding measure, but here we use online newspaper articles as data sample. The third measure tries to capture perceived economic policy uncertainty by the relative share of monthly Google search frequencies connected with policy uncertainty in Norway. The fourth measure aims to capture perceived firm-specific policy uncertainty by applying textual analysis to company annual reports. We do, however, not control for other likely related measures when estimating firm exposure to policy uncertainty, such as economic uncertainty in general.

The motivation for investigating the role of economic policy uncertainty in the cross-section of returns may be anchored in Merton's (1973) model foundation of the Intertemporal Capital Asset Pricing Model (ICAPM). Merton states that investors want to hedge against future stochastic shifts in consumption and set of investment opportunities as it includes information about investors' marginal utility of wealth. Policy uncertainty likely implies such a shift in investors' investment opportunity sets based on the empirical research of businesses acting more carefully by reducing employment, investment (Baker et al., 2016) and dividend payments

(Walkup, 2016) when policy uncertainty increases. Furthermore, the variable is difficult to diversify against due to its omnipresent nature.

We contribute to literature in two main ways. First, we try to capture EPU by online articles and Google searches and create a firm-specific measure of policy uncertainty in the Norwegian market. By doing this, we hope to contribute to the debate concerning how to capture investor attention as well as increasing the robustness of our analysis. Second, we investigate the implications of policy uncertainty in the cross-section of Norwegian stock returns. To the best of our knowledge, we are the first to do this.

The rest of the thesis is organized as follows. Section 2 provides a review of existing literature, including the debate on ways of capturing policy uncertainty and former research on the relationship between risk and expected returns, both internationally and in Norway. Section 3 describes the empirical methods applied to our study, both for how we capture policy uncertainty and for the estimation of risk premia. Section 4 describes the data samples and adjustments made. We present our findings and their implications when testing our hypothesis in Section 5. Section 6 presents the conclusion of the paper, including limitations and future research possibilities.

2 Literature Review

This section presents relevant literature. The first subsection presents theory related to measures of economic policy uncertainty while the second subsection presents asset pricing literature.

2.1 An Introduction to Policy Uncertainty

This part of the literature review briefly introduces the term economic policy uncertainty before describing its implementations in research.

2.1.1 Measuring Uncertainty

Researchers have called attention to uncertainty in the financial world ever since the *The Age of Uncertainty* by Galbraith (1977). Still, it took several years before its effect on financial markets were studied. A universal definition of uncertainty is not agreed upon in literature, however, there is no question behind its importance. This study focuses on uncertainty stemming from governments and its interaction with stock returns. Al-Thaqeb and Algharabali (2019) define policy uncertainty as: "the economic risk associated with undefined future government policies and regulatory frameworks". Throughout the paper, the terms economic policy uncertainty (EPU) and policy uncertainty are used interchangeably.

In existing literature, there is no doubt of governments' impact on the business environment (Friedman, 1968; Rodrik, 1991). Van Den Bosch (1994) advocates that government should be included as an independent determinant of competitive advantage. Multiple studies, such as Colak et al. (2017), Jens (2017) and Pastor and Veronesi (2012) indicate that firms tend to act more carefully when facing high economic policy uncertainty. The studies find that the cost of capital increases when EPU is high, resulting in firms taking less part in investments (Gulen & Ion, 2013), reduce employment (Baker et al., 2016), set in motion little capital raising (Colak et al., 2017), decrease M&A activity (Bonaime et al., 2018) as well as reducing capital paid out to their equity investors (Walkup, 2016).

As a universal way of defining policy uncertainty is not agreed upon, neither is a method of capturing it. However, a measure established as the standard for quantifying policy uncertainty was introduced by Baker et al. (2013). They introduce a measure of EPU by combining three different components; a news-based component, a component based on federal tax codes set to expire and a component concerning disagreement among economic forecasters related to policy variables. The news-based measure is given the most weight in their index and is the methodology used in this thesis. It is quantified by extracting monthly numbers of newspaper print articles related to EPU divided by the total amount of articles published. An article is classified as concerning EPU if it includes words related to (i) the economy, (ii) government institutions and (iii) uncertainty. Baker et al. (2016) expand their study, finding reduced investment rates and decreasing employment growth when levels of EPU increases.

The uncovering of policy uncertainty's effect on businesses led to innovation in measurement techniques. Azqueta-Gavaldòn et al. (2020) apply machine learning to news articles. The approach is based on a continuous selection of words inserted into an unsupervised machine learning algorithm to pick up articles related to EPU. The strength of the machine learning method arises from its ability to split EPU into specified topics, allowing different economic responses to be connected with distinct elements of policy uncertainty. Although utilizing a different methodology than Baker et al. (2013), the authors find that the measures correlate strongly. The machine learning process introduces an element of bias due to selection of topics based on human judgement. Consequently, the selection may include words capturing other phenomena not necessarily related to policy uncertainty because it is based on picking the most prominent words within each topic.

There is an ongoing debate concerning which data sources reflect investor attention in the most optimal way. Policy uncertainty is comprised of several unobservable variables and may not be objectively captured by any specific methodology. While the traditional methodologies are based on print media, a growing number of studies implement online sources as data foundation. Da et al. (2011) find that an index created from Google searches captures investor attention in a more timely manner than other measures, and provide evidence that their estimate in fact captures investor attention. Several researchers leverage these findings by creating Google search-based measures applied to economic policy uncertainty. Examples are Bontempi et al. (2016), Castelnuovo and Tran (2017) and Donadelli (2015). The two first studies utilize a comprehensive list of words, while the latter includes fewer words to create a proxy for policy uncertainty. While some introduce Google search frequencies as a way of capturing uncertainty, others suggest more subtle changes to the method of Baker et al. (2013). Kim (2020) advocates the use of online articles instead of print-based ones when measuring investor attention grounded on the findings that online news articles have a stronger impact on asset prices. The author links this to the tone of news, but the findings could also stem from other sources, like online news being more available to the public.

While aggregate measures of policy uncertainty have received the most attention in literature, some aim to capture the phenomena at the firm level (Hassan et al., 2019; Nie et al., 2020). Hassan et al. (2019) use a method of computational linguistics on transcripts from conference calls in conjunction with earnings releases to measure firm level political risk in the U.S. The authors investigate language patterns that correlate with policy topics by training their model on political texts, and allows it to recognize the association with political risk by including synonyms of risk and uncertainty. They find firms cut back on hiring and investment when exposed to increasing political risk. Nie et al. (2020) utilize text mining tools to measure firm-specific policy uncertainty perceived by Chinese firms. They measure firm level policy uncertainty by conducting sentence analysis on annual reports. The measure is quantified by looking at sentences including at least one word related to uncertainty and government policy in relation to total number of sentences. The authors find, in line with Hassan et al. (2019), that increases in perceived firm level policy uncertainty leads to decreased investment and increased holding of financial assets. Based on the literature, we introduce four methods for capturing EPU; (i) a print article-based, (ii) an online article-based, (iii) a Google Trends-based measure and (iv) a firm-specific measure derived from annual reports, of which will be fully elaborated in Section 3.

This thesis contributes to existing literature by proposing new ways to measure policy uncertainty in Norway, a field of research which has received limited attention previously. To the best of our knowledge, measures of economic policy uncertainty using Google search frequencies, online articles and textual analysis applied to company information has not been employed before in Norway. By introducing new methods of capturing policy uncertainty, we hope to add new perspectives on how to capture investor attention in the Norwegian market.

2.2 Asset Pricing

> *"Price is expected discounted payoff. This fundamental relation underlies all asset pricing. The discount factor is an index of `bad times`. Because investors are willing to pay more for assets that do well in bad times, the risk premium on any asset is determined by how it covaries with the discount factor." (Cochrane & Culp, 2003)*

Every approach to asset pricing builds on the principle included in the quote above, stating that the price of an asset should equal the present value of future expected cash flows. The authors connect this to consumption smoothing and risk aversion; the phenomena that investors have a concave utility function, which implies diminishing marginal utility of consumption, and therefore care about consumption smoothing. The theorem states that investors need a reward to carry systematic risk. This basis of investor behaviour is essential in asset pricing models and their connection to risk premia. In this section, we present past and current ideas of how to explain variation in stock returns using factor models.

2.2.1 Factor Pricing Models

Cochrane (2000) states that the consumption-based model is the very foundation of asset pricing. Other theories like the Capital Asset Pricing Model (CAPM), Intertemporal Asset Pricing Model (ICAPM) and Arbitrage Pricing Theory (APT) are specializations of the consumption-based model created due to its unsatisfactory empirical performance. The Capital

Asset Pricing Model (CAPM) by Sharpe (1964), Treynor (1961), Lintner (1965) and Mossin (1966) try to answer how investment risk should affect expected returns. The model is built on the proposal that not all types of risk should influence asset prices. It uses portfolio theory by Markowitz (1952) to argue that diversifiable risk should not carry a risk premium. The CAPM illustrates a linear relationship between systematic risk and expected returns

$$E(R^i) = R^f + \beta * E(R(m) - R^f), \tag{2.1}$$

where R^i and $R(m)$ are the expected returns of asset i and the market portfolio m respectively, and β is a measure of the volatility of asset i compared to the market. The theory states that all investors will adjust their portfolios by maximizing Sharpe ratio until stock prices alter to equilibrium so that CAPM holds. However, the CAPM relies on assumptions that will not likely hold in the real world, such as investors being able to borrow and lend at the risk-free rate.

In the wake of the CAPM limitations, others formed linear factor models to better capture equity risk premiums. Two of the most renowned ones are the ICAPM and the APT. The ICAPM was introduced by Merton (1973) and includes investor wealth as a state variable, making the model consider lifetime consumption decisions. The main contribution of the ICAPM is the supplementary state variables which includes investors' desire to hedge against future consumption shortages or changes in the investment opportunity set. Merton (1973) argues that a constant opportunity set of investment is unrealistic and claims that it is state-dependent, leading investors to change portfolio composition. The ICAPM states that expected returns are a linear function of the risk-free asset, the market portfolio and a third asset; the portfolio hedging against changes in the set of investment opportunities.

Ross' (1976) Arbitrage Pricing Theory is another alternative to the empirically inaccurate CAPM. The APT introduces a framework that describes expected returns of assets as a linear function of the asset's risk concerning a factor set representing systematic risk. Ross (1976) argues that the linear function between expected returns and factor loadings holds if equilibrium prices offer no arbitrage opportunities. The APT has its advantages compared to the CAPM as it relies on fewer assumptions, while at the same time allowing for more than one factor to explain expected returns. The assumptions include: (i) asset returns can be explained by systematic factors, (ii) investors can diversify away risk by constructing portfolios and (iii) properly diversified portfolios have no possibility of arbitrage. Given that investors hold diversified portfolios, exposure to idiosyncratic risk will be voided and investors will only be exposed to systematic risk. If there are no arbitrage opportunities and the real systematic risk factors are known, assets with the same exposure to systematic risk factors must have equal expected returns. This is derived from the law of one price. Even though the CAPM and the APT may seem similar, the theoretical foundations of the models vary considerably. The CAPM is an equilibrium model while the APT is supported by a no-arbitrage premise. The APT may be expressed as

$$E(R^i) = X\lambda = R^f + \beta\lambda , \tag{2.2}$$

where β is a matrix of sensitivities of asset return i to risk factors and λ is the risk premium associated with the different risk factors.

Even though the APT is convenient by being easier on assumptions, it is challenging to use because it does not specify which systematic risk factors it should include to describe expected returns. The precedent way of dictating which factors to include is through empirical research of company specific attributes as substitutes for systematic risk factors. However, as the APT introduced a framework that allows for several factors when explaining expected returns, it paved the way for models based on the same foundation. These can in many ways be interpreted as variations of Ross' model. The models include the renowned Fama-French Three- (Fama & French, 1993) and Five-factor (Fama & French, 2015) models and the Carhart Four-factor model (Carhart, 1997). The models are based on anomalies which the CAPM is not able to capture. Studying these anomalies have attracted significant attention in financial research. Within finance, an anomaly describes a pattern of deviations of real returns from what is expected in financial models. One well described anomaly is the "small-cap" effect. Banz (1981) and Reinganum (1981) find a negative relationship between size and returns by discovering that companies of smaller market capitalization are consistently associated with higher returns. An explanation for this effect is offered by Klein and Bawa (1977). They argue that amount of company information is positively correlated with firm size. If sufficient information is not available, investors will demand a risk premium to hold smaller firms due to uncertainties associated with lack of information. Another well investigated anomaly is the "Book-to-Market" effect. It is based on the rationale that firms with relatively high book value of equity compared to market capitalization offer fundamentally cheaper equity. The effect is well documented by Fama and French (1992), Basu (1977) and Lakonishok et al. (1994). Basu

(1977) finds this effect by using a P/E ratio while Fama and French (1992) makes use of the ratio between book value and market capitalization[1].

In their Three-factor model, Fama and French (1993) include their empirical findings of value stocks outperforming growth stocks and small-cap stocks outperforming large-cap stocks, in addition to the market factor. A few years later, Carhart (1997) expanded the model by including a momentum factor based on the discovery of returns correlating with prior returns. Many have tried to develop these models further, with Fama and French's evolution of their own Three-factor model being one of the most renowned ones; The Five-factor model (Fama & French, 2015). The theoretical reasoning for adding the new factors were based on the dividend discount model with the assumptions of Miller and Modigliani (1961), stating that book-to-market ratios, expected investment and expected profitability are linked to expected returns of stocks. Consequently, each of these factors should absorb all variation in stock returns when controlling for the other two (Fama & French, 2006). Thus, the model was expanded with an investment and a profitability factor. The profitability factor was developed from the rationale that, holding all else equal, higher profitability should lead to higher expected stock returns. The positive relationship between profitability measures and expected returns has been empirically verified by papers such as Haugen and Baker (1996), Novy-Marx (2010) and Fama and French (2015) on U.S. stock returns and by Nichol and Dowling (2014) on stock returns from the UK. Furthermore, the intuition for including the investment factor is that for constant levels of profitability and book-to-market ratio, an increase in assets by investing is associated with lower expected returns. Several explanations with foundations in behavioural economics are offered to explain the negative relationship between investment and expected returns, such as the overinvestment hypothesis introduced by Stulz (1990). The negative link is proven empirically in U.S. stock returns by Aharoni et al. (2012) and Fama and French (2015). Fama and French (2017) extends the geographical scope of their study, finding that a Five-factor model allows for absorption of additional patterns in average returns when adding European and Asian Pacific stocks to the study.

Leveraging the assumption of globally integrated financial markets, findings from the U.S. stock market should hold across geographical markets. However, this has proven not to be the case. Studies find that a factor model applying to all markets is difficult to come by, and that

[1] Fama and French finds the P/E measure to be redundant in multivariate regressions.

regional variations often outperform global counterparts (Fama & French, 2012; Griffin, 2001). Several researchers document imperfections that offer reasoning for this. Dumas and Solnik (1995) find support for the existence of foreign exchange risk premia, meaning that stock returns in different markets price exchange rate risk derived from distinct markets. Transferring this to the Norwegian stock market, Sæbø (2008) and Næs et al. (2009) find that the size and market factors are highly significant for explaining returns. However, the findings of Næs et al. (2009) regarding the book-to-market factor using a simple sorting method is more ambiguous and less systematic as they only find this effect significant in two out of three sub-periods between 1980 and 2006. Furthermore, the authors find very limited support for the momentum effect in the Norwegian stock market. Apart from the research of Sæbø (2008) and Næs et al. (2009), little documentation regarding systematic risk premia in the Norwegian stock market is published. The two factors most recently added to the Fama and French Five-factor model, operating profitability and investment, is yet to be assessed in a published study covering the Norwegian market. However, the factors' ability to explain returns have been investigated in some master theses' which have found them to not add any explanatory power relative to the Three-factor model (Hoel & Mix, 2016; Bakken, 2019). However, as the evidence against the factors are limited and they have proven to be useful controls in international markets, we include them as a specification in our analysis. Nevertheless, the fact that they have not been found significant in studies conducted at the OSE is something to keep in mind when evaluating if sensitivity to economic policy uncertainty may explain variation in Norwegian stock returns.

2.2.2 Macroeconomic Variables

Several studies on pricing of macroeconomic factors in cross-sectional stock returns have been performed (Bali et al., 2017; Brogaard & Detzel, 2014). The motivation is anchored in Merton's model foundation of the ICAPM regarding investors' desire to hedge against future stochastic shifts in consumption and set of investment opportunities, and that these variables may include information about investor's marginal utility of wealth (Merton, 1973). Consequently, state variables correlating with alterations in consumption and investment opportunities should be priced in the equity premia of stock returns. Næs et al. (2009) investigates the properties of the oil price in relation to stock returns but find that the variable is not a priced risk factor in Norway. Brogaard and Detzel (2014) argue that economic policy uncertainty is a variable that affects investment opportunities by its forecasting effect on stock market returns. Furthermore, they find evidence of EPU obtaining a significant negative risk premium when explaining stock returns in the U.S. The authors argue that this is because increases in policy uncertainty portray a worsening in investment opportunities and that investors want to hold stocks which hedge against this. In other words, investors desire stocks which returns covary positively with levels of policy uncertainty. The investment opportunity set of an individual includes all the investments the investor is capable of in a time period. Increases in policy uncertainty may worsen the investment opportunity set as it is found to be associated with reduced employment growth and decreased dividend payments (Baker et al. 2016). Holding assets that negatively covary with levels of policy uncertainty may amplify volatility of consumption, which investors want to avoid. Furthermore, Brogaard and Detzel (2014) argue that EPU contains relevant information distinct from general economic uncertainty on the basis of the Pastor and Veronesi (2012) model. This thesis expands the research of Brogaard and Detzel (2014) by taking a regional view of EPU as a factor premium and utilizing various measures of policy uncertainty. Keeping in mind that different anomalies exist in different markets, it is interesting to investigate if investors' required rates of stock returns vary by assets' sensitivity to EPU in the Norwegian stock market.

This thesis contributes to existing literature by investigating the role of policy uncertainty in the cross-section of stock returns in Norway. To the best of our knowledge, the role of policy uncertainty in the cross-section of Norwegian stock returns is not covered in literature. Furthermore, we increase the robustness of our analysis by using several distinct measures of economic policy uncertainty. The thesis aims to increase attention to the research field of macro variables and their impact on firms in Norway.

3 Methodology

This section aims to present an in-depth description of the methods applied in the thesis. We split this into two subsections; (i) methodology covering our measures of capturing policy uncertainty and (ii) the methods applied to investigate the role of economic policy uncertainty in the cross-section of returns.

3.1 Creating EPU Indices

Four different techniques are applied to capture policy uncertainty. These include (i) a print newspaper-based index, (ii) an online newspaper-based index, (iii) a Google search-based index and (iv) a firm-specific measure. The reason for employing four different measures is to increase the robustness of our analysis because there is no universally accepted method of capturing policy uncertainty. By introducing new ways of quantifying policy uncertainty to the Norwegian market we hope to expand this field of research.

3.1.1 Geographical Scope

To measure if investors at the OSE require a systematic risk premium for holding stocks sensitive to policy uncertainty, it is important that our indices reflect the economic policy uncertainty perceived by the marginal investor at the OSE. Given that Norwegians own more than 60% of capital at OSE (Oslo Børs, 2020a), this is likely to be the most important investor group. Note that this implies that nearly 40% of capital at OSE comes from elsewhere. This could be controlled for by capturing the policy uncertainty foreign investors believe is coherent with the Norwegian market and weigh the measures by relevance. We keep this in mind, but do not perform this exercise due to the time-consuming nature of it, combined with lack of access to such data in other geographic regions. Given that the majority of capital at the OSE is owned by Norwegians, we believe that EPU measures grounded in Norwegian sources is a suitable proxy for the policy uncertainty inherent in the marginal investor at the OSE. From this reasoning, we utilize only Norwegian newspapers and Google searches conducted in Norway.

3.1.2 EPU Based on Newspaper Articles

We follow the method of Baker et al. (2016) when creating newspaper-based indices. The method involves creating a frequency of EPU articles relative to the total number of articles published. An article must contain at least one word within each of three categories to be classified as an article related to EPU. This includes one synonym of the word “economy”, one word related to “governmental policy” and one synonym of the word “uncertainty”. We implement these criteria in Atekst Retriever as it allows for multiple conditions by utilizing the conjunctions “AND” and “OR”. The full list of words is illustrated in Table A.1.

For newspapers to represent a suitable reflection of policy uncertainty, a necessary presumption is that they are capable of capturing public perception of uncertainty without manipulating it. If news articles preceded policy uncertainty by manipulating public perception, we would have an issue with our indices being leading. Hopkins et al. (2017) finds that newspapers in the U.S. do not precede public perceptions of the economy, but that media coverage rather reflects public perception. Given that economic policy is closely related to the economy, there may be reason to believe that newspapers also have the capabilities to reflect such subjects. Given that these assumptions are transferable to Norway, our newspaper-based measures is likely to be a good proxy for policy uncertainty.

As noted in the literature review, mainly two methods have been established for extracting indices based on newspaper articles; keyword search-based methods and machine learning processes. A machine learning process may introduce substantial bias due to selection of topics based on human judgement, while the method of Baker et al. (2013) should exclude articles not related to EPU by using a search criterion including three separate word categories. In this study, we follow the methodology of Baker et al. (2013) when utilizing newspapers, both for print and online articles. This allows us to incorporate a search criterion that has been subject to extensive auditing, and consequently provide trustworthy results. Baker et al. (2016) perform extensive human auditing of newspapers, finding that their search-based measure has a correlation of 0.93 with the index created by manually classifying articles. This emphasizes the accuracy of this method. The bias of picking up articles not actually concerning EPU should thus be limited given that these findings are transferrable to Norwegian newspaper articles.

Furthermore, the selection of which newspapers to include is of high importance in the pursuit of a trustworthy index. One issue could arise from newspapers having their own agendas. If the newspapers in our sample had a political agenda, this could alter our newspaper indices based on conditions such as head of government. This could lead to larger focus on policy uncertainty in times where other political parties than those affiliated with the newspaper's views were in power. DellaVigna and Hermle (2014) finds, even though investigating this in movie reviews, that media reputation is an important factor for preventing biased coverage. If this is the case for other parts of news coverage, we have reason to believe that newspaper reputation is a powerful disciplining force for unbiased coverage. With this in mind, we only include reputable national newspapers. For the print-based index, we include Aftenposten, VG and Dagbladet. For our online-based measure, we select DN.no, VG.no and Dagbladet.no. The newspapers included are not identical due to Atekst Retriever differing somewhat in regard to what sources it keeps for online and print news. For print, it does not have access to DN, while it includes Aftenposten which is not included for online articles. However, we keep as many reputable national newspapers as possible in order to capture EPU in the most representable way. We construct normalized time series of each newspaper by standardizing the series to a unit standard deviation and then assigning them a mean value of 100. This is so that the indices may be compared to the Google-based index as well as other indices. At last, we weigh the elements equally to obtain the indices. An equal weight is employed because we believe the newspapers are equivalently important sources for capturing policy uncertainty, and we want to keep the variation of all inputs. Although these newspapers are large in Norwegian scale, some are substantially larger, such as VG, and we believe we get a more sensitive index by equally emphasizing the variation of the different sources.

3.1.3 EPU Based on Google Searches

We also introduce a measure of capturing EPU based on Google Trends since the historical search frequencies should capture investor attention in an objective, direct manner. Studies like the one performed by Da et al. (2011) have indicated that this is the case. With close to 90% market share, Google is the ideal source when measuring investor attention from online searches. There are several ways of utilizing Google Trends to capture attention as Google reports search frequencies for search terms and topics. Search terms are the specific words used in a search, while search topics will include all terms related to the topic. By using search topics, one may capture a lot of noise because the search frequencies pick up related topics that are not necessarily connected with policy uncertainty. Furthermore, the method entails less transparency as the user is not inclined to a full overview of what the different topics may or may not reflect at each point in time. Based on these grounds, we choose to construct the Google-based EPU measure from search terms.

We build our Google EPU measure based on the method of Donadelli (2015). As Donadelli (2015) estimates policy uncertainty from frequencies of the search terms “US stock market”, “US politics” and “US Fed”, we obtain our index by including the three equivalent terms in Norwegian; “Oslo Børs”, “Norsk politikk” and “Norges Bank”. The terms are included in a single query in Google Trends in order to be weighted together and are thus ready to use. This is because Google Trends normalize search data by scaling each search term relatively, assigning their peak period to a score of 100. As for our other measures, we adjust the time series to a unit standard deviation and assign a mean of 100 to ease comparison of indices. Note that when only including a few keywords, issues concerning biases may arise as these in reality may be searched more frequently during phenomena of attention distinct from policy uncertainty. This is particularly true for the term “Oslo Børs”, which search frequency is likely to vary with various economic events. However, while we keep this issue in mind, we accept the index as the best possible proxy for policy uncertainty due to Google Trends having issues with the Norwegian language when broadening our scope to longer lists of keywords.

We initially wanted to construct our Google search index following the structure of papers like Bontempi et al. (2016) and Castelnuovo and Tran (2017), which use a comprehensive list of specific keywords. While the former weighs all words equally, the latter split them into different categories where the individual categories are weighted. Although these papers aim to estimate other types of uncertainty, one could follow the same structure when capturing policy uncertainty, and adopt the categories stated as most important from Baker et al. (2013); taxes, spending, monetary and regulatory policy. The methods may allow for greater precisions when measuring policy uncertainty as only terms strictly related to governmental policy would be included. However, Google Trends require a minimum volume of searches within a time period to report search frequencies. When making an index based on a comprehensive list of words, e.g. “styringsrente”, we are not able to obtain consistent historical frequencies of search volumes due to limited data for these particular words. By implementing our alternative approach, we are able to retrieve consistent estimates.

3.1.4 Firm-specific EPU: Textual Analysis

While the majority of studies have investigated the role of policy uncertainty in the cross-section of returns by using aggregate measures of policy uncertainty, Hassan et al. (2019) applies the idea of measuring firm exposure to EPU more directly. While the authors use the obtained firm-specific measure of EPU to forecast variables such as investment, we aim to use a similar measure to explain expected stock returns in Norway. The intention is that by analysing documents produced by firms, one may obtain a more unmediated measure of their perceived uncertainty and sensitivity towards government policy matters. The method is based on the assumption that firms more uncertain and sensitive towards future policy shocks will mention terms related to this topic more frequently than other firms. We gather inspiration from papers aiming to measure firm-specific EPU such as Hassan et al. (2019) and Nie et al. (2020), however, applying a distinct methodology. While Hassan et al. (2019) apply textual analysis to analyst earnings calls, we utilize annual reports. Furthermore, we use a document-term matrix to obtain our firm-specific measure of EPU. The method involves a mathematical matrix describing the rate of occurrence of terms in a collection of documents. Using this method, we extract the number of times terms in our policy uncertainty related dictionary are mentioned in a set of texts. The matrix is constructed so that each column represents a specific term, and each row represents a document. This allows us to analyse the development in usage of different words for specific firms over time. Before the analysis, we process our documents by removing signs and uninformative words, such as prepositions, pronouns and numbers. Then, we implement a dictionary of words related to policy uncertainty inspired from the list of words we apply to newspaper articles, as presented in Table A.2. The estimate of a specific firm's uncertainty about future policy shocks is then calculated as the number of words associated with government policy relative to the total amount of words in the annual report. We standardize the series of each firm to a unit standard deviation. When measuring firm-specific levels of policy uncertainty through textual analysis, we do not apply the criteria that terms need to be mentioned in relation to words concerning the economy or uncertainty. First, since annual reports by nature deal with affairs of economic nature, use of words related to government policy should be associated with circumstances related to the economy. Second, under the assumption that relative word frequency reflects perceived uncertainty regarding future business conditions, the measure should be a suitable estimate for perceived policy uncertainty.

We use English versions of annual reports as the standard when measuring firm-specific EPU. However, some companies only publish their annual reports in Norwegian. To keep a satisfying amount of data, we accept this. We control for this by including both Norwegian and English words related to government policy in our dictionary, so that our algorithm is able to deal with both languages. For each term considered to be related to policy uncertainty, we include one word for each language. As a resulting effect, it may be the case that an English term is more natural to use than its Norwegian counterpart, thus resulting in more hits. However, as we do not compare levels of EPU between firms, but rather how this measure correlate with stock returns for one firm at a time, the possible alteration should not affect our analysis. Furthermore, some firms may be more inclined to use words related to governmental policy in their reports, thus constantly obtaining higher estimates. Again, since we are concerned with correlation between the measure of EPU and stock returns, we are more interested in variation, and consequently this is not an issue for the analysis.

3.1.5 Comparing the Indices

We find that our newspaper-based indices covary significantly with a correlation of 0.88. From this, it seems like the online-based and print-based measure primarily capture the same variation. The Google search-based measure behave a little differently, having a correlation with the print article-based measure of 0.60. The indices spike very similarly in 2008 during the financial crisis and in March 2020 during the outbreak of the Covid-19 pandemic in Europe. However, we do not see the same fluctuations in the periods in between, where our newspaper-based measures fluctuate more. This is particularly true during the oil price crisis. An explanation for this could be related to potential biases associated with the Google index as previously discussed, but it could also be the case that these are periods where Norwegian investors in fact have perceived governmental policy as less uncertain. The correlation between our aggregate indices is visualized in Table 3.1 and the indices are visualized in Figure 3.1. Additionally, in a global world, it is useful to understand whether our regional measures of EPU capture any region-specific variation compared to existing global measures. When comparing our news-based measures of policy uncertainty for Norway to a global measure, we obtain a correlation of approximately 74%. This implies that our indices do capture country-specific events. The global measure is obtained from the official website concerning economic policy uncertainty (Baker at al., 2020).

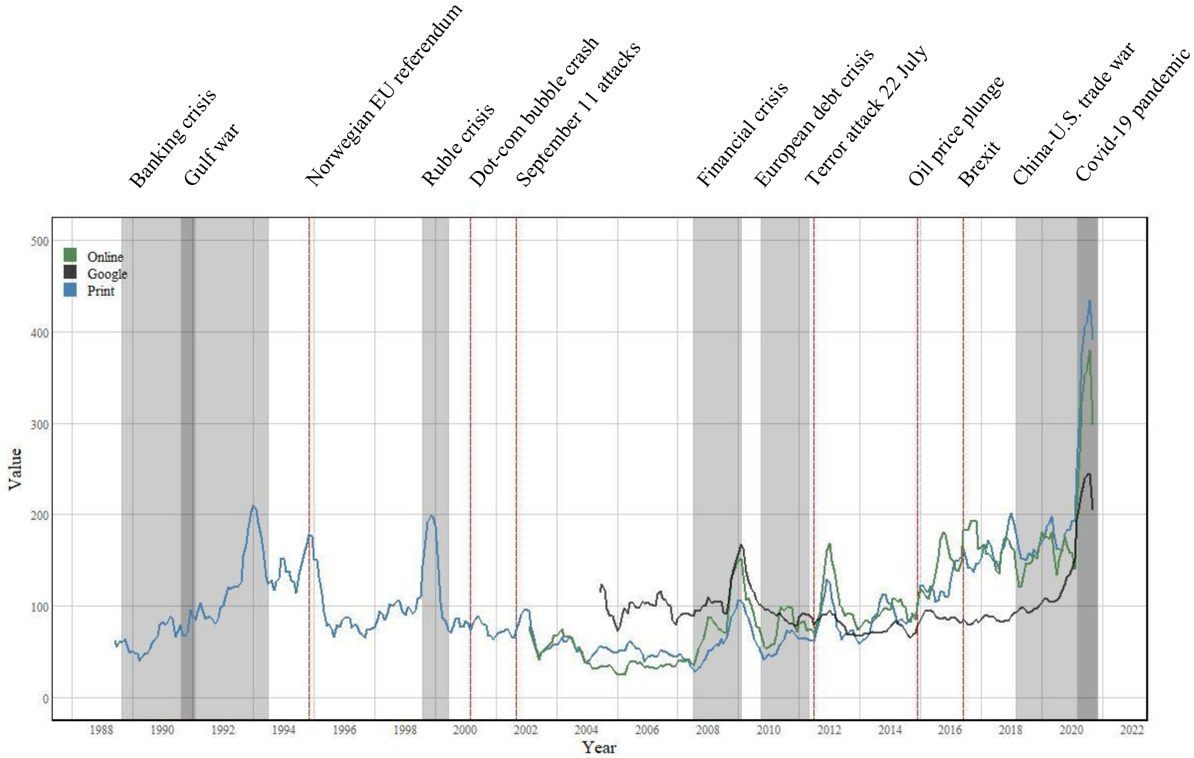
Table 3.1: Correlation Matrix

Correlation between aggregate measures of economic policy uncertainty.

	Print articles	Online articles	Google searches
Print articles	1.00	0.88	0.60
Online articles	0.88	1.00	0.57
Google searches	0.60	0.57	1.00

Figure 3.1: EPU Indices

Display of aggregate measures of EPU based on (i) print newspaper articles, (ii) online newspaper articles and (iii) Google search frequencies. Red lines mark relatively brief incidents while grey shading marks more extensive events. All indices displayed based on six months rolling average.



3.2 Factor Pricing

Expected returns of assets ought to be a function of their exposures to factors correlating with expected consumption in the future. This section describes the steps taken to investigate the role of policy uncertainty in the cross-section of Norwegian stock returns and the motivation for applying factor models.

3.2.1 Model Specifiations and Factor Construction

Our study aims to expand existing literature by estimating if policy uncertainty carries a systematic risk premium in Norway by applying the APT framework. To ensure that the factor does not capture variation already picked up by other factors, the estimates will be obtained while controlling for multiple factor models. We control for market, size, value, investment and operating profitability by applying the CAPM, Fama-French Three-factor model and Fama-French Five-factor model when investigating if policy uncertainty explains variation in stock returns. The CAPM and Fama-French Three-factor models are chosen based on existing factor model research on Norwegian stock returns with Næs et al. (2009) finding the market, size and book-to-market factors (between 1980-2000) to be priced in Norwegian stock returns. The Fama-French Five-factor is added to the analysis due to its increased explanatory power in explaining international stock returns relative to the other two models, even though this has not yet been investigated in a published study covering the Norwegian market. Furthermore, we construct the control factors from accounting and stock data rather than retrieving them from public sources. The factor estimates rely on the assumptions and data used to form them and we want the factors to be consistent with our data sample.

3.2.2 Testing Framework

An established method of estimating parameters for asset pricing models is the Fama-Macbeth framework. In short, the model approximates the exposures (β) and risk premiums for any given risk factor of which one may argue is connected with determining asset prices. The method allows for using panel data by first estimating assets' exposures to certain factors through time before using these exposures to estimate whether the given factors are systematic or not. The estimation of factor exposures (β) for each asset is performed through linear regression, which can be expressed as

$$R_t^i - R_t^f = \alpha_i + \sum_n \beta_{i,n} f_{t,n} + \epsilon \ , \tag{3.1}$$

where R_t^i is the return of asset *i* at time *t*, R_t^f is the risk-free rate at time *t*, α_i is a constant which in theory should equal zero, $\beta_{i,n}$ is the exposure of asset *i* to factor *n* and $f_{t,n}$ is the excess return associated with factor *n* at time *t*. The factor realization, $f_{t,n}$, is estimated using mimicking portfolios. A mimicking portfolio is a tradeable combination of assets which equal the exposure of the wanted underlying asset. The method for constructing our mimicking portfolios is described in Section 3.2.4.

After estimating factor exposures from time series regressions, a second regression will estimate whether the factors are systematic. This is conducted using cross-sectional regressions of the form

$$R_t^i - R_t^f = \alpha_{n,t} + \sum_n \lambda_{n,t} \beta_{i,n} + \epsilon \ , \tag{3.2}$$

where the left-hand side of the equation represent the excess return of asset *i* in month *t*, $\beta_{i,n}$ is the exposure to a given factor as described above and $\lambda_{n,t}$ is the factor premium for asset *n* at period *t*. We estimate out-of-sample because the stock exposures are retrieved from firm characteristics of the former period to avoid look-ahead bias. The cross-sectional regressions obtain risk premia estimates for each factor in every month. Since the Fama-Macbeth methodology prohibits risk premia varying over time, the realized value of a risk premium associated with a factor is calculated as

$$\hat{\lambda}_n = \frac{1}{T}\sum_{t=1}^{T}\lambda_{n,t} \ , \tag{3.3}$$

where T is the total number of time periods and $\lambda_{n,t}$ is the risk premium associated with factor *n* at time *t*. The estimated risk premia are assumed to be drawn from a normal distribution so that our time invariant estimator $\hat{\lambda}_n$ is unbiased. When deciding if the factor carries a systematic risk premium, t-tests are conducted using the sample mean and standard deviation. However, this method may lead to econometric issues (Ødegaard, 2020b). The applicable bias is noted as errors-in-variables. The issue may arise because the exposures of each asset (β) are first estimated in the time series regressions and then applied to the cross-sectional regressions.

3.2.3 Describing the Factors

When estimating the explanatory power of policy uncertainty in Norwegian stock returns, it is useful with an introduction to the control variables and the EPU variable. These variables are proxies for firm characteristics connected with risk premia. The factor mimicking risk associated with size is estimated using the market capitalization of a stock, which includes the value of all shares outstanding, noted as

$$Market\ capitalization_t = Share\ price_t * \ Shares\ outstanding_t. \tag{3.4}$$

Furthermore, the value characteristic is estimated as a firm's book value of equity relative to its market capitalization

$$B/M_t = \frac{Book\ value\ of\ equity_{t-1}}{Market\ capitalization_{t-1}}, \tag{3.5}$$

where Fama and French (1993) define book value of equity as

$$
\begin{aligned}
&Book\ value\ of\ equity_{t-1} \\
&\qquad = Stockholder\ equity_{t-1} + Deferred\ taxes_{t-1} \\
&\qquad + Investment\ tax\ credit_{t-1}.
\end{aligned}
$$

Following Davis et al. (2000), we use the difference between total assets and total liabilities as a substitute for book value of equity when stockholder equity is not available. The investment factor aims to capture a firm's investment behaviour by using asset growth as a proxy, defined as

$$INV_t = \frac{Total\ assets_{t-1}}{Total\ assets_{t-2}} - 1. \tag{3.6}$$

The final characteristic added in Fama and French's framework is the factor aiming to capture operating profitability traits, which Fama and French (2015) define as

$$OP_t = \frac{Total\ revenue_{t-1} - Total\ operating\ expenses_{t-1} - Interest\ expenses_{t-1}}{Book\ equity_{t-1}}. \tag{3.7}$$

We estimate a firm's sensitivity to policy uncertainty by regressing excess returns on inventions of the applicable EPU index using the expression

$$R_{t+1}^{i} - R_{t+1}^{f} = \alpha_{i,t} + \sum_{i} \beta_{i,t} EPU_t + \epsilon\ , \tag{3.8}$$

where $\beta_{i,t}$ is the stock's estimated sensitivity to policy uncertainty at time *t* and EPU_t is the value of innovations in the respective measure of policy uncertainty at time *t*. We define innovations as the relative change of the relevant index between time *t* and *t-1*. A rolling regression with a window of 36 observations is applied to estimate stock exposure to EPU. We include the criterion that stocks need a minimum of 18 return observations over the past 36-month period in order to obtain estimates. The rolling window is applied as firms may change operations over a medium to long period of time, and exposure to policy uncertainty may change correspondingly. Given that policy uncertainty may covary with other types of uncertainty, like economic uncertainty in general, we would ideally include proxies for these as control variables in our regressions to obtain unbiased estimates. The NOVIX, which is a volatility index based on the VIX methodology to reflect uncertainty in the Norwegian stock market, would be a suitable control variable. However, the NOVIX index starts in April 2016, and its time span is therefore too limited for our analysis. Since we aim to address policy uncertainty in the Norwegian market, we consider global measures as inconsistent with our analysis. Therefore, we do not include other uncertainty measures as control variables when estimating exposure to policy uncertainty, although we keep this potential issue in mind.

3.2.4 Mimicking Portfolios

To determine if various factors explain stock returns, we need a method of estimating the realization of different factors at each time, noted as $f_{t,n}$ in equation 3.1. Factor mimicking portfolios are helpful instruments for this purpose. We create the mimicking portfolios with backward looking characteristics in the end of June each year as companies in Norway are required to make last year's annual report publicly available by the first of July[2]. The mimicking portfolios are constructed based on the specific factors which are motivated to explain the shortcomings of the CAPM at OSE. We form our factor mimicking portfolios based on two size groups and three groups of the other factors. The motivation for sorting the mimicking portfolios on respectively two and three characteristics at a time is that Fama and French (2015) find that other classifications does not perform significantly better. Thus, the portfolios are constructed from a breakpoint at 50% for the size characteristic and combined with groups constructed using the 30[th] and the 70[th] percentiles as breakpoints for the other factors. The portfolios are meant to isolate the respective firm characteristics by using double sorts. This method allows for the realization of our factors. SMB (Small Minus Big) aims to capture the size characteristic by retrieving returns when an investor is long a diversified portfolio of the smallest stocks and short a diversified portfolio of the largest stock. With our method using 2 x 3 sorts, this is done by buying the three smallest portfolios and selling the three largest portfolios for each sorting method. The return of the size mimicking portfolio at any given time in context of the Five-factor model is thus calculated as the average of three double sort methods:

$$SMB_{B/M} = {}^{1}\!/_{3}\,(Small\ Value + Small\ Neutral + Small\ Growth) - {}^{1}\!/_{3}\,(Big\ Value + Big\ Neutral + Big\ Growth)$$

$$SMB_{OP} = {}^{1}\!/_{3}\,(Small\ Robust + Small\ Neutral + Small\ Weak) - {}^{1}\!/_{3}\,(Big\ Robust + Big\ Neutral + Big\ Weak) \tag{3.9}$$

---

[2] The portfolios mimicking the control variables are based on financial characteristics published in annual reports.

$$SMB_{INV} = {}^{1}\!/_{3}\,(Small\ Conservative + \ Small\ Neutral + Small\ Aggressive) - {}^{1}\!/_{3}\,(Big\ Conservative + Big\ Neutral + Big\ Aggressive)$$

$$\text{SMB} = {}^{1}\!/_{3}\left(SMB_{B/M} + \ SMB_{OP} + SMB_{INV}\right).$$

Note that stocks will be assigned to the same size groups, so a simple sorting method would produce identical values for the SMB factor. Since we use 2x3 sorts with size as basis, the other factor returns are calculated by buying two portfolios and selling two portfolios. The HML (High Minus Low) factor describes the book-to-market effect by expressing the return of being long a diversified portfolio of high book-to-market stocks and short a portfolio of low book-to-market stocks, expressed as

$$HML = {}^{1}\!/_{2}\,(Small\ Value + Big\ Value) - {}^{1}\!/_{2}\,(Small\ Growth + Big\ Growth). \tag{3.10}$$

The CMA (Conservative Minus Aggressive) factor expresses the investment behaviour effect by reflecting the return of a portfolio consisting of the most conservative firms minus a diversified portfolio of the most aggressive firms, and is calculated as

$$CMA = \ {}^{1}\!/_{2}\,(Small\ Conservative + Big\ Conservative) - {}^{1}\!/_{2}\,(Small\ Aggressive + Big\ Aggressive). \tag{3.11}$$

The RMW (Robust Minus Weak) factor aims to represent the operating profitability effect by describing the return of holding a portfolio of the most profitable firms and selling a diversified portfolio of the least profitable firms, derived as

$$RMW = \ {}^{1}\!/_{2}\,(Small\ Robust + Big\ Robust) - {}^{1}\!/_{2}\,(Small\ Weak + Big\ Weak). \tag{3.12}$$

At last, the EPU mimicking portfolio, NMP (Negative Minus Positive), aims to capture the policy uncertainty effect by buying the stocks with the lowest EPU exposure (i.e. most negative covariance) and selling the stocks with the highest exposure to EPU (i.e. most positive covariance), so that

$$NMP = {}^{1}\!/_{2}\,(Small\ Negative + Big\ Negative) - {}^{1}\!/_{2}\,(Small\ Positive + Big\ Positive). \tag{3.13}$$

In theory, investors want to hold stocks that covary positively with EPU, thus the realization should be a positive estimate in order to be theoretically sound. Finally, we construct the market factor as the return of a value weighted portfolio of the majority of stocks at OSE in excess of the monthly risk-free rate. These variables and their origins are described in Section 4. The EPU mimicking portfolios are constructed at the same time period as the financial factors to ensure consistency[3].

3.2.5 Test Portfolios

We follow the methodology of Jensen et al. (1972) and Fama and Macbeth (1973) by using portfolios as opposed to single stocks as explanatory variables when conducting the analysis. The authors argue that employing portfolios helps reduce idiosyncratic risk and therefore generate better factor loading estimates and risk premia estimates accordingly. When testing if policy uncertainty carries a systematic risk premium in Norwegian stock returns while controlling for the Fama and French factors, we would ideally sort our portfolios by all factors at the same time to achieve full isolation. That is, controlling for all other factors believed to affect returns in order to obtain unbiased estimates. However, as this would imply sorting at n different dimensions, where n is the number of factors, we would obtain an excessive number of portfolios compared to our sample of stocks. This is because the Norwegian stock market is a relatively limited stock market in terms of number of listings. Ødegaard (2020c) argues that a diversified portfolio should consist of at least 10 stocks. Hence, we apply the method of double sorting to our test portfolios. This is conducted by separating stocks into three distinct groups for each individual factor characteristic before forming the portfolios based on the size attribute as well as one of the remaining factors. Thus, the size characteristic is used as the basis for all sorts. The test portfolios are created at the end of June each year to represent implementable trading strategies. We obtain nine Size-B/M portfolios, nine Size-OP portfolios and nine Size-INV portfolios. We split the stocks into three quantiles based on each firm characteristic because of the relatively limited sample size. Our test portfolios satisfy Ødegaard's criteria for classification as diversified portfolios for the most part. However, there are some issues particularly related to portfolios characterized as large-cap value stocks, large-cap stocks with weak operating profitability, small-cap stocks with robust operating profitability and large-cap stocks with aggressive investment behaviour. Furthermore, we note that our method of merely

[3] The portfolios mimicking financial factors relies on annual reports and consequently cannot be formed more frequently.

double sorting may bias our estimates due to factors potentially correlating. However, since we want to conduct our analysis using diversified portfolios to reduce idiosyncratic risk, we do not possess the luxury of being able to sort more granularly. The average number of stocks per test portfolio is presented in Table A.3.

3.2.6 Alternative Methods of Estimating Risk Premia

Applying the Fama-Macbeth framework to asset pricing entails many advantages, such as control for time effects, intuitive interpretation of results and suitable treatment of unbalanced panel data such as ours. Furthermore, papers have found the Fama-Macbeth method to produce consistent estimates with reliable t-statistics and to be more efficient than generalised least squares estimates when using long time series (Skoulakis, 2008). Nonetheless, several papers have addressed the potential issue of procedures such as Fama-Macbeth producing incorrect standard errors (Petersen, 2007; Pagan, 1984). Cochrane (2000) reason that pooled time series and cross-section OLS in most finance applications may produce standard errors that are off by a factor of 10. However, Cochrane (2000) proclaims that this is mainly an issue for corporate finance applications, and not a problem to the same extent for asset pricing estimates due to returns being close to independent. Still, Cochrane (2000) argues that even though the risk premia estimates from the Fama-Macbeth procedure are unbiased, their standard errors are not. However, our analysis is not sensitive to this issue, as we do not use the standard errors of the point estimates, but merely the standard errors of the cross-sectional regression estimates to obtain sampling errors. These are not linked to the standard errors of premia estimates at each time *t*. Petersen (2007), who investigate the phenomena of biased standard errors in a collection of applications, concludes that the performance of different methods of estimation relies on the structure of the data sample, and that advice regarding how to deal with this is relatively limited in literature. Furthermore, he shows that the Fama-Macbeth method deals well with time effects in the residuals, producing unbiased standard errors, but that this is not the case when the data includes firm effects. Therefore, our standard errors may be biased, ultimately affecting our t-statistics. Following this, Petersen (2007) and others claim that it is useful to check if results are the same using different methods of calculating standard errors, to verify the robustness of the chosen approach. Given the time-consuming nature of estimating various measures of policy uncertainty, combined with the assurances of Cochrane (2000), we avoid other methodologies of estimating the risk premia in this paper such as the Generalized Method of

Moments introduced by Hansen (1982), and perform our analysis solely using the Fama-Macbeth method.

3.2.7 Theoretical Motivation for Using Factor Models

This section presents the theoretical motivation for applying factor models to our analysis. First, we present consumption smoothing theory before including the concept in an asset pricing setting.

Utility Functions and Risk Aversion

The expected utility hypothesis describes preferences of which an economic decision maker is concerned with. The theory estimates the likely utilities of different outcomes and suggests the rational decision maker will choose the option with the highest expected utility by maximizing

$$E[u(x)] = p_1 * u(x_1) + p_2 * u(x_2) + \cdots + p_n * u(x_n)\,, \tag{3.14}$$

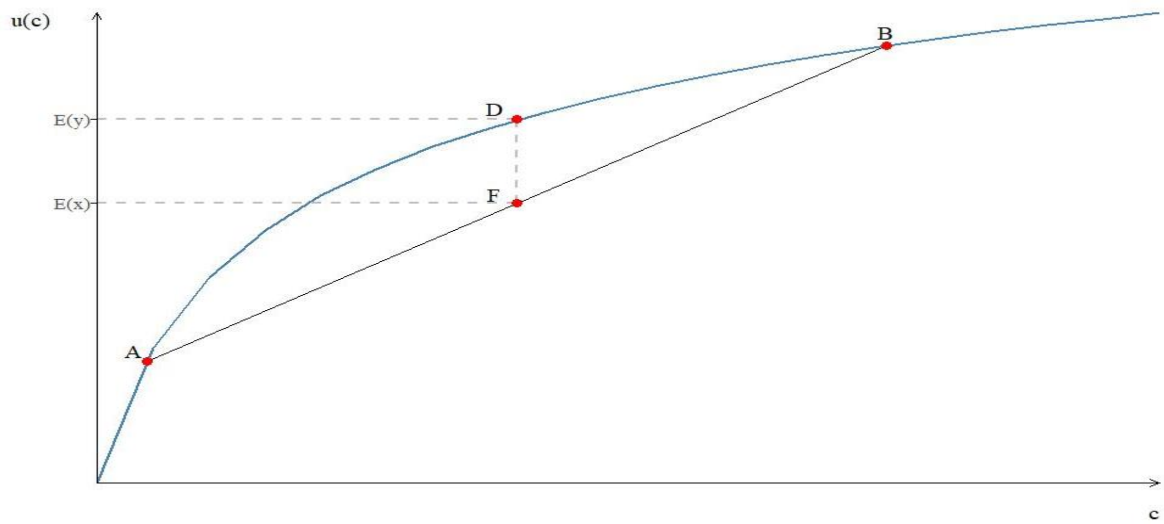
where p_i is the probability of outcome i and $u(x_i)$ is the utility of outcome x_i. This equation introduces the idea that the option with the highest expected value or consumption is not necessarily the option that grants the highest expected utility, but that this depends on the decision maker's valuation of the options. The Von Neumann-Morgenstern (Prokop, 2014) utility theorem introduces a term to quantify the subjective value of these potential outcomes. One version of a Von Neumann-Morgenstern utility function with the assumption of constant relative risk aversion can be written as

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}\,, \tag{3.15}$$

where c_t is a value of an outcome at time t, $u(c_t)$ is the utility of the given outcome and γ is a constant specifying the degree of relative risk aversion. From the expression we gather that a higher γ is associated with higher risk aversion. Drawing the function with $\gamma > 0$ we see that it is represented as a concave utility function, illustrated in Figure 3.2. The role of risk premia in the case of uncertainty considering a risk averse investor with diminishing marginal rate of utility can be illustrated in the following scenario. The decision maker gets the option between choosing to receive a certain sum of money D or to receive a financial asset that can be worth A in one state and B in another state. The states are revealed the second the decision maker receives the asset, and the expected value of the asset $\frac{(A+B)}{2} = F$, so that the expected values of the outcomes are the same. Furthermore, the expected utility of choosing the certain sum of money D can be noted as $E(y)$ while the expected utility of choosing the financial asset with outcomes A or B can be noted as $E(x) = 0.5 * u(A) + 0.5 * u(B)$. However, given the concave utility function of the decision maker, his expected utility $E(y) > E(x)$, so that the decision maker will require some compensation in order to choose the financial asset. This compensation takes place as a risk premium, which determines how different asset prices are formed. The example illustrates that investors do not like uncertainty about consumption, as stated by Cochrane (2000). Because economic agents want to smooth their consumption, they value assets that perform well in bad times higher than assets performing well in good times. Since the intervention of government is difficult to diversify away and is a variable likely to affect investment opportunities, consumption smoothing theory lays the very foundation of this thesis as we explore risk premia connected with policy uncertainty.

Figure 3.2: Utility Function

Utility function $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ illustrated with γ = 1.2.



Asset Pricing

Factor pricing theory tries to explain how asset prices and risk premiums are formed based on consumption smoothing theory. Investors have a utility function which is increasing, but at a decreasing rate. This reflects the constant desire to consume more while marginal utility is diminishing; we value another unit of consumption more when we have less resources. These behavioural patterns give us the intuition to grasp how risk premiums are formed. Since investors want to smooth their consumption, and value certainty as explained in the section above, they will need a risk premium to hold assets with high betas. Cochrane (2000) introduces a general framework of asset pricing by advocating a discount factor view. In the framework, we want to quantify the value of any flow of uncertain cash flows. Noting that an investment is concerning the choice of trading consumption today for consumption in the future, an investor needs to find the value at time t of a payoff in the next period, x_{t+1}. If this is a stock, the payoff at time *t+1* will include $x_{t+1} = p_{t+1} + d_{t+1}$, where p_{t+1} and d_{t+1} is the price and dividends in the period *t+1*. x_{t+1} is a random variable, making the investor uncertain of how much he will receive in advance. However, he can estimate the probability of different outcomes in line with the expected utility hypothesis expressed in equation 3.14. We can formalize this investor trade-off as

$$U(c_t, c_{t+1}) = u(c_t) + \beta E_t[u(c_{t+1})] , \tag{3.16}$$

where c_t is the consumption at time *t*. To find the utility of the outcomes, $u(c_t)$, we can use a power utility function of the form represented in equation 3.15. This formalization lets us correct for risk and delay of cash flows by capturing the investors' impatience and risk aversion. The β in the expression captures willingness to postpone consumption and is considered as the subjective discount factor, while the γ in the utility function captures risk aversion. Introducing a trading mechanism to the framework, where an investor can buy and sell as much as he pleases of the payoff x_{t+1} at a price p_t, we can note his problem as

$$\begin{aligned} \max_{\xi} \mathrm{u}(c_t) + E_t\, \beta \mathrm{u}(c_{t+1})\ s.t. \\ c_t = e_t - p_t\xi \\ c_{t+1} = e_{t+1} + x_{t+1+\xi} , \end{aligned} \tag{3.17}$$

where e_t is the original consumption level without investing and the amount of assets he chooses to buy has notation ξ. Inserting the two constraints into the objective function, differentiating and setting it equal to zero gives us the following first-order condition for the investor's maximization problem

$$p_t u'(c_t) = E_t[\beta u'(c_{t+1}) x_{t+1}]. \tag{3.18}$$

This is the standard marginal condition for an optimum where the left side of the equation describes the utility loss the investor bears by purchasing another unit of the asset while the right-hand side quantifies the increase in utility by acquiring one payoff at time *t+1*. The investor will trade the asset until equilibrium. Here, the marginal utility obtained from one extra payoff in the future is equal to the marginal utility lost by buying another unit; the investor is indifferent between consumption today and tomorrow. Rearranging this expression, we get the central asset-pricing formula

$$p_t = E_t\left[\beta \, \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}\right]. \tag{3.19}$$

The first-order condition reveals the price p_t of which the investor is willing to buy an asset with an uncertain payoff x_{t+1}. This is decided by his impatience β and his risk aversion, indirectly drawn from the relationship between the utility of consuming a certain amount today compared to an uncertain amount tomorrow. Cochrane introduces a convenient notation for the stochastic discount factor

$$m = \beta \, \frac{u'(c_{t+1})}{u'(c_t)} \,, \tag{3.20}$$

So that by substitution the asset-pricing formula can be rewritten as

$$p_t = E_t(m_{t+1} x_{t+1}) \tag{3.21}$$

The expression states that variations in prices and returns have three origins. The first one is due to time variation in the stochastic discount factor *m*, the second is from various shocks impacting the stochastic discount factor and the third is shocks to expected cash flows, *x*. This implies that all variables affecting the marginal utility of wealth for investors have the potential to be priced as factors. Introducing the special case where the payoff is a return and the price is 1, we get

$$1 = E\left(mR^i\right) \tag{3.22}$$

Given that the asset pricing model states that returns can vary, but expected discounted returns should always equal 1 and including a decomposition of the covariance we obtain

$$1 = E(m)E\left(R^i\right) + cov\left(m, R^i\right) \tag{3.23}$$

Following that $R^f = \frac{1}{E(m)}$, we find

$$E\left(R^i\right) - R^f = -R^f cov\left(m, R^i\right), \tag{3.24}$$

stating that all assets must give a risk adjusted return in addition to the risk-free rate. It follows that investors will demand a risk premium for assets performing poorly when times are bad because holding these assets is expected to make the investor's consumption level more volatile. Given that we can rewrite equation 3.24 as

$$E(R^i) = R^f + \left(\frac{cov(R^i, m)}{var(m)}\right)\left(-\frac{var(m)}{E(m)}\right), \tag{3.25}$$

and by defining $\beta_{i,m}$ as $\left(\frac{cov(R^i,m)}{var(m)}\right)$ and λ_m as $\left(-\frac{var(m)}{E(m)}\right)$, we get an intuitive expression for a simple beta factor model

$$E(R^i) = R^f + \beta_{i,m}\lambda_m , \tag{3.26}$$

where β_m is the assets regression coefficient of the return R^i on the stochastic discount factor and λ_m can be interpreted as some universal premium for loading risk for all assets. The expression simply states that the expected return of an asset is the quantity of risk in that asset times the price of risk. Factor pricing models build on this foundation in the setting of systematic risk and is consequently the motivation behind choice of methodology in our analysis. In the setting of a multi-factor model in the APT framework, we are aiming to estimate the quantity of policy uncertainty risk carried by assets in order to obtain estimates of the price of this risk.

4 Data

This section presents the data samples and describes adjustments made.

Time Period

We conduct our analysis on the longest time period possible. The Compustat Global database provides accounting data for companies listed at OSE from 1989 and stock data from 1986. Since we need accounting data from the two previous years to estimate the investment factor and we use a 36-month rolling regression window to estimate firm sensitivities to policy uncertainty, we are able to investigate the period between December 1992 and June 2019. However, three of our EPU measures are obtained at a shorter time horizon and will thus limit our analysis to a shorter time frame.

4.1 EPU Data

4.1.1 Newspaper-based Indices

We obtain data from Atekst Retriever for both of the newspaper-based indices. The Atekst Retriever database allows the user to employ conditions when searching for articles. Furthermore, it includes a vast selection of both online-and print-based sources and there are no duplicates, neither for print articles nor for online articles. Given that we want to investigate the longest time period possible, we use the first data available from 1988. It is based on VG and Aftenposten for the first years before also including Dagbladet in 1996. Our online article-based measure starts in September 2001 and is based on VG.no, Dagbladet.no and DN.no in the whole period. There is some variation in newspapers included in the two indices due to data availability from Atekst Retriever, where the online version of Aftenposten and the print version of DN are not available. Descriptive statistics for the newspaper sources used in the newspaper-based indices are illustrated in Table 4.1.

Table 4.1: Summary Statistics of Newspaper Sources

Newspaper	Type	First article	EPU articles	Total articles	Avg. monthly EPU articles
VG	Print	01/88	1,813	1,094,829	4.6
Dagbladet	Print	01/96	1,589	873,665	5.4
Aftenposten	Print	01/88	7,999	1,980,870	20.4
VG.no	Online	09/01	1,324	593,218	5.8
Dagbladet.no	Online	09/01	1,485	449,037	6.5
DN.no	Online	09/01	3,919	337,026	17.1

4.1.2 Google-based Index

The data used to construct our Google-based index is retrieved from Google Trends. This is Google’s publicly available database for historical search frequencies. The database is a convenient tool for measuring attention to various search terms as it publishes relative search frequencies for terms in any given region and time period. The frequencies are adjusted relatively from previous search history and should therefore be suitable for measuring developments in attention. Google publish search frequencies from January 2004. Our index starts from January 2005 as this is when our search terms start to obtain consistent values. We extract monthly search frequencies for constructing our Google-based measure of EPU to make it coincide with the periodicity of returns. Furthermore, this is the most granular frequency available for longer time horizons.

4.1.3 Firm-specific EPU

The annual reports used to obtain a measure of firm-specific policy uncertainty are retrieved from the respective companies' homepage and Newsweb. Newsweb is a search engine operated by the OSE which include all company announcements of listed firms. We prefer using company homepages when obtaining annual reports. If reports are unavailable, we retrieve the rest from Newsweb. We gather the annual reports of firms with at least six years of reported stock returns in the period 2009-2019 and apply textual analysis to the reports which have corresponding returns. To avoid look-ahead bias, the report corresponding to a given month of stock returns will be the newest one publicly available for the respective company. Given that Oslo Stock Exchange demand listed firms to publish annual reports by the 1st of July in the following year, we use this as a proxy for publicity date for all companies. That is, a report in year t is used to explain the returns in the second half of year $t+1$ and the first half of year $t+2$. Consequently, we form implementable trading strategies for the period July 2009 until June 2019 by retrieving annual reports from 2008 until 2017. We utilize a total of 723 annual reports from 82 different companies. Summary statistics are presented in Table 4.2.

Table 4.2: Summary Statistics of Annual Reports

Development in annual reports. Number of annual reports vary by number of listings at OSE.

Year	No. of companies	Mean EPU terms	Mean terms[4]	Mean frequency[5]
2008	65	22	22,802	0.083
2009	73	21	22,337	0.083
2010	76	21	22,961	0.074
2011	79	21	22,986	0.078
2012	78	21	23,132	0.076
2013	78	22	23,252	0.082
2014	74	26	24,549	0.089
2015	72	24	24,773	0.078
2016	65	28	25,246	0.086
2017	63	30	26,710	0.092

[4] Terms excluding uninformative words, e.g. prepositions and pronouns.

[5] Frequency is calculated as $\frac{Number\ of\ EPU\ terms}{Total\ number\ of\ words}$ for each annual report.

4.2 Stock Data

Stock data are retrieved from the Compustat Global database. This database provides both stock price data and accounting data needed to construct the factors. Compustat Global provides accounting data at a yearly frequency, while the database provides daily data for stock prices in Norway. Since we follow the established practice of utilizing monthly returns when assessing factor risk premia, these are reconstructed to describe monthly returns. After creating the factors and cleaning the raw data, our data frames are merged. Consequently, our final data frame does only include observations with both stock returns and accounting data. Summary statistics for the stock data are presented in Table 4.3 and the steps for adjusting the data are illustrated in Table. 4.4. The resulting number of firms per year in our sample after the cleaning and merging process is described in Figure 4.1, and the return distribution of our final stock sample is illustrated in Figure 4.2. Note that there is a relatively limited number of companies listed at OSE, which could be a limiting factor for the analysis.

Figure 4.1: Number of Companies in Final Sample

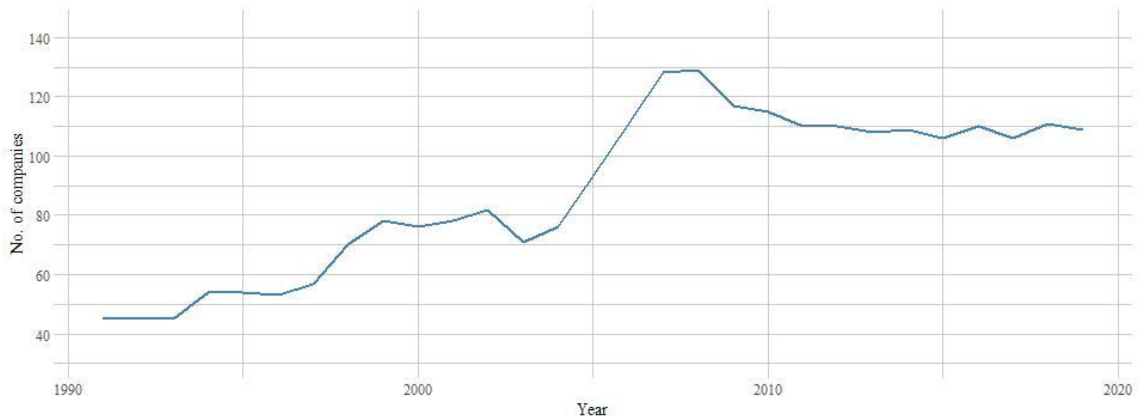


Table 4.3: Summary Statistics of Stock Data and Risk-free Rate

Summary statistics of stock data and monthly risk-free rate from adjusted sample.

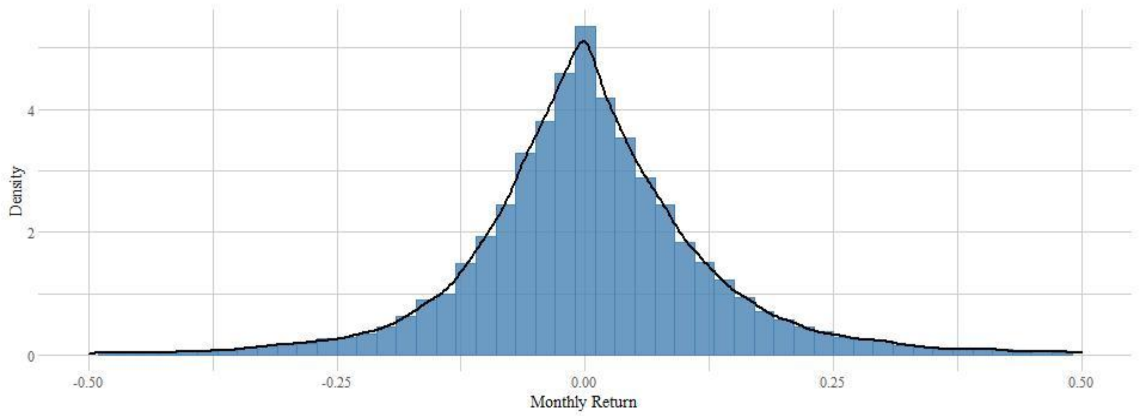
	N	Mean	SD	Min	Max
Monthly returns	27,900	0.008	0.142	-0.904	3.580
Risk-free rate	27,900	0.003	0.002	0.001	0.021
Market cap (NOK)	27,900	10,090,283,462	43,172,379,787	2,540,373	766,222,749,760
Shares outstanding	27,900	145,980,491	396,208,371	752,475	4,103,777,581

4.2.1 Stock Types and Exchanges

Following Bali et al. (2017), we only include common shares in our analysis. This group will include shares of different classes, meaning that they have slight modifications, e.g. additional voting rights. These stocks are usually classified as A or B shares. In Norway, there is no established practice regarding which modifications belong to which stock class, and it is therefore difficult to separate these incidents in an organized manner. However, even with slight alterations to stock properties, these shares behave like equity stock, and as Compustat Global defines these as common shares, they are all included in our analysis. Furthermore, our data sample include dual listed firms, meaning firms are also listed at a foreign exchange. These stocks are merely duplicates of the local listings, often to increase share liquidity. Following this, all observations listed at foreign exchanges are removed. This also eliminates the potential issue concerning stocks only being listed at a foreign exchange. We want to estimate sensitivity to policy uncertainty in Norway and these stocks may be subject to distinct policy environments. Ultimately, we restrict our analysis to stocks listed at OSE and exclude over-the-counter (OTC) stocks as these are not traded on regulated exchanges.

Figure 4.2: Return Distribution

Return distribution in the final data sample after all adjustments.

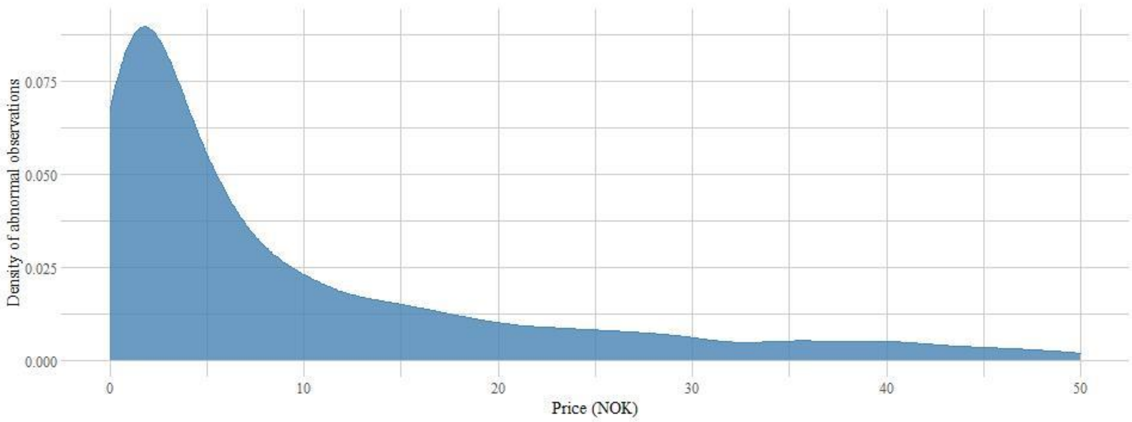


4.2.2 Penny Stocks

In asset pricing literature, it is widespread to remove stocks of very low value, so called penny stocks. This is because penny stocks may misrepresent the analysis as they often contain issues related to illiquidity and inflated returns. Consequently, even slight fluctuations in prices may lead to considerable changes in returns. In Figure 4.3, this issue is illustrated by showing the return distribution of stocks by share value in our sample. In Norway, there is no practical definition of what is considered as penny stocks. Ødegaard (2020c) suggests defining penny stocks as stocks trading below 10 NOK per share as well as stocks having a total market capitalization of less than 1 MNOK. He argues that all observations should be removed in a year when a stock meet one of these criteria. We follow Ødegaard's restriction regarding market capitalization by removing all observations in a year where the market capitalization has been below 1 MNOK, but we modify Ødegaard's recommendation regarding share price to keep a satisfying number of observations in our sample. Therefore, we follow the directive of the OSE. It requires all stocks to trade at more than 1 NOK to be listed (Oslo Børs, 2020b). However, it allows companies to trade below 1 NOK for a period of up to 6 months. Thus, we propose eliminating all observations in a year where a share has traded below 1 NOK in any month. By excluding the full year, we also omit observations where stocks trade over 1 NOK. We argue this is reasonable given that stocks trading below 1 NOK in a time period is likely affected by microstructural issues that may also be present in surrounding months. This method of omitting penny stocks allows us to control for the most extreme cases of microstructure issues, while at the same time keeping a satisfying sample size.

Figure 4.3: Distribution of Abnormal Returns by Share Value

Abnormal returns defined as the 2.5% highest and lowest returns each month.



4.2.3 Financial Firms

The majority of financial firms have higher levels of leverage than non-financial firms due to the nature of their operations. Whereas considerable leverage may imply distress for non-financial firms, this may not be the case for financial firms. The different leverage structure of financial firms may distort our method of factor creation, and consequently our analysis. Thus, we follow the method of Fama and French (1992, 1993) and discard financial firms from our analysis.

4.2.4 Negative Book Value of Equity

A few stocks in our sample are reported with negative book value of equity. This can only happen if the value of the firm's liabilities exceeds the value of its assets, meaning that the book equity is negative. However, this has no meaningful economical interpretation as shareholders' liabilities are limited. Since two of our factors, book-to-market and operating profitability, are constructed using book value of equity, the factors may lose their economic interpretation for these assets. As an example, a firm with negative book value of equity and negative profitability will be perceived by our operating profitability factor as a profitable company. From this reasoning we follow the common practice in academics by omitting observations with negative book value of equity (Brown et al., 2007; Fama & French, 1995). Furthermore, we also omit stocks with reported zero asset values, as these observations are likely to be subject to misreporting.

4.2.5 Calculation of Returns

Since Compustat Global provides daily stock data for Norway, we need to calculate monthly returns. This is done by using prices of the last day each month. Since some stocks do not have returns from the last day of the month, but for days close to it, we slightly ease this criterion by using the last observation reported at or after the 25th day of each month. This allows for keeping more observations. An implication of this modification is that returns may be calculated over slightly different time periods. However, we consider the effect to be limited. As returns are affected by events such as dividends and stock splits, we need to adjust for this. The Compustat database provides such a tool through its adjustment factor which allows for uncomplicated accommodation. Employing this, we calculate returns as

$$r_t^i = \frac{\frac{p_t^i}{adj\,fac_t^i} - \frac{p_{t-1}^i}{adj\,fac_{t-1}^i}}{\frac{p_{t-1}^i}{adj\,fac_{t-1}^i}} \tag{4.1}$$

where r_t^i is the return of asset i in month t, $\frac{p_t^i}{adj\,fac_t^i}$ is the adjusted share price of asset i at time t and $\frac{p_{t-1}^i}{adj\,fac_{t-1}^i}$ is the adjusted price of asset i in the previous period. As mentioned, the resulting return distribution of our sample is illustrated in Figure 4.2.

4.2.6 Exchange Rates

All stock observations from Compustat Global are reported in NOK. However, this is not the case for the accounting data. Some companies have their main operations in countries other than Norway, thus utilizing a reporting currency different from NOK. In our sample, these include EUR and USD. To obtain all data in a consistent manner, we exchange all values to NOK. This is achieved by downloading monthly exchange rates for EUR/NOK and USD/NOK from Norges Bank. Note that there is one company registered with EUR as reporting currency in 1998. Given that the EUR currency was introduced 1st of January 1999, we omit this observation.

4.2.7 Risk-free Rate and Market Returns

We retrieve estimates of the risk-free rate in Norway from professor Bernt Arne Ødegaard's website (Ødegaard, 2020a). The risk-free rate is a forward-looking estimate for borrowing at a monthly basis. The market index is constructed from our sample retrieved from Compustat Global which includes the returns of the majority of stocks at Oslo Stock Exchange. The smallest stocks and financial firms are omitted as stated earlier, but no stocks are removed due to accounting issues, like negative book value of equity. We form the returns this way to create a market index that is representative for the OSE. The index is value weighted.

Table 4.4: Data Adjustments

Description of adjustment steps and impact on sample size.

Accounting data		
	Observations	Difference
Compustat Global	5,690	
Omit financial firms	4,460	-1230
Omit foreign exchanges	4,148	-312
Omit duplicates	4,141	-7
Omit zero assets	4,117	-24
Omit book equity < 0	3,995	-122
Calculate INV and omit NAs	3,400	-595
Stock data		
	Observations	Difference
Compustat Global	1,352,246	
Omit preferred stock	1,334,353	-17,893
Omit financial firms	1,088,833	-245,520
Omit foreign exchanges	890,640	-198,193
Calculate monthly returns	44,101	-846,539
Omit penny stocks	42,607	-1,494
Merge file with accounting data	27,900	-14,707

5 Results and Discussion

The results of our analysis and their implications are presented in this section. First, we give an overview of the EPU factor by illustrating spreads on portfolios sorted by policy uncertainty exposure. Thereafter, we present our findings following the steps of the Fama-Macbeth framework; factor realizations, results from time series regressions and results from cross-sectional regressions. To control for various firm-specific effects, we include control models in our regressions; the CAPM, the Fama-French Three-factor model and the Fama-French Five-factor model. However, as this paper aims to investigate the role of policy uncertainty as risk premium in the Norwegian stock market, we focus on the findings related to the EPU factor. Ultimately, we investigate if EPU carries a negative systematic risk premium in line with the theoretical foundation when controlling for various well renowned multi-factor models.

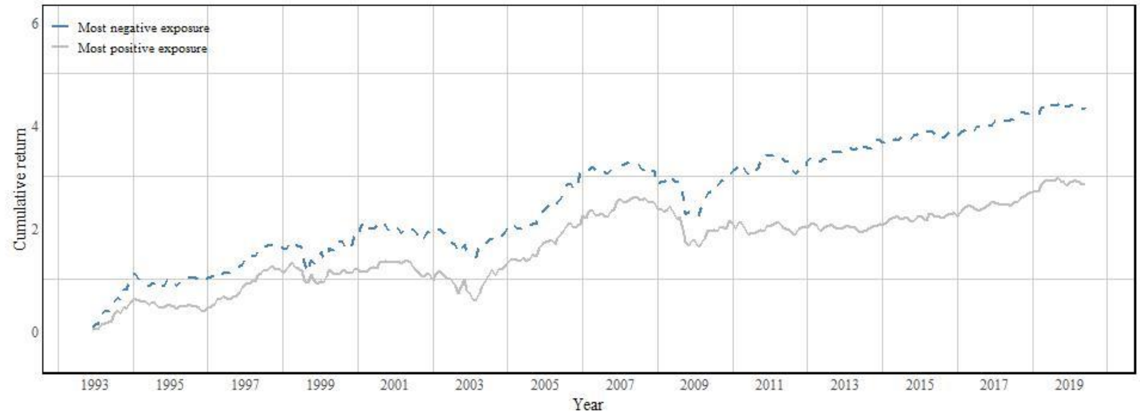
5.1 Overview of the EPU Factor

We start the analysis by investigating how average returns vary with exposure to EPU to obtain an overview of the role of policy uncertainty in the Norwegian market. This is performed by simply sorting the stocks by exposure to policy uncertainty in the end of June each year and plotting the cumulated returns of the portfolio including the stocks with the highest positive exposure to EPU against the portfolio consisting of the stocks with the lowest (highest negative) exposure to EPU. The plots give an initial overview of EPU as a potential risk factor as the graphs are based on simple sorts. Tables with portfolio means and p-values are included to determine if the spreads are significant.

Cumulative returns for the extreme portfolios when measuring EPU by print articles is illustrated in Figure 5.1. We observe a slight tendency that the portfolio with the most negative exposure to policy uncertainty outperforms the portfolio with the most positive exposure to EPU over time. Nevertheless, the probability of observing returns this different, given the true mean is similar, is 43%. Consequently, we cannot reject the null hypothesis and claim that a significant spread exists. The results are seemingly the same when inspecting Figure 5.2, where EPU is measured by online articles. The cumulative returns of the two extreme portfolios follow each other closely, with the most positive exposure stocks having slightly higher returns until 2008, before the stocks most negatively exposed to EPU obtain moderately higher returns in subsequent years. The corresponding table states that the average returns are approximately the same over the whole time period. With a p-value of 99%, there is no case for advocating that the returns are different. The story is the same when measuring policy uncertainty by Google search frequencies in Figure 5.3. The portfolios obtain nearly equal cumulative returns during the sample period, and a p-value of 91% states no evidence against the null hypothesis. At last, we observe the cumulative return spread between the extreme portfolios measured by firm-specific EPU in Figure 5.4. The situation is the same as for the other measures, even though the p-value is smaller at 18.9%. However, this is far from our 5% confidence level and we cannot infer that the populations are different. Nevertheless, it is interesting to investigate if policy uncertainty does affect expected stock returns when controlling for other factors because the portfolios may carry different loadings on other risk factors.

Figure 5.1: EPU Portfolio Spreads – Print Newspaper Articles

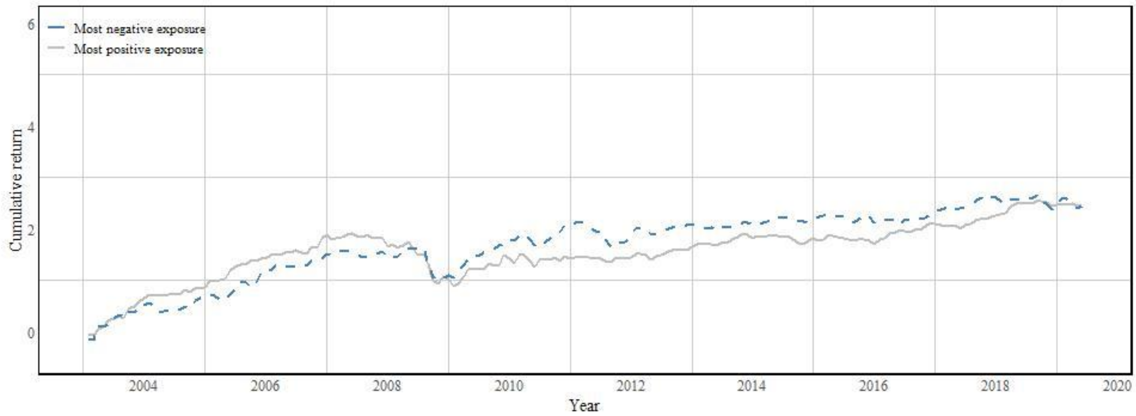
Cumulative excess returns of the two extreme EPU portfolios in the period of 1992-2019. Portfolios are simple sorted by exposure to EPU measured by print articles. The portfolios include the companies with the 1/3 most negative EPU exposure and the 1/3 most positive exposure. Portfolios are constructed at the end of June each year with weights held until next June. The associated t-test determines if there is a statistically significant difference in mean of returns between the two groups, with the null hypothesis stating that no such difference exists. A low p-value indicates evidence to reject the null hypothesis.



Mean most negative	Mean most positive	t-statistic	p-value
0.014	0.009	0.79	0.43

Figure 5.2: EPU Portfolio Spreads – Online Newspaper Articles

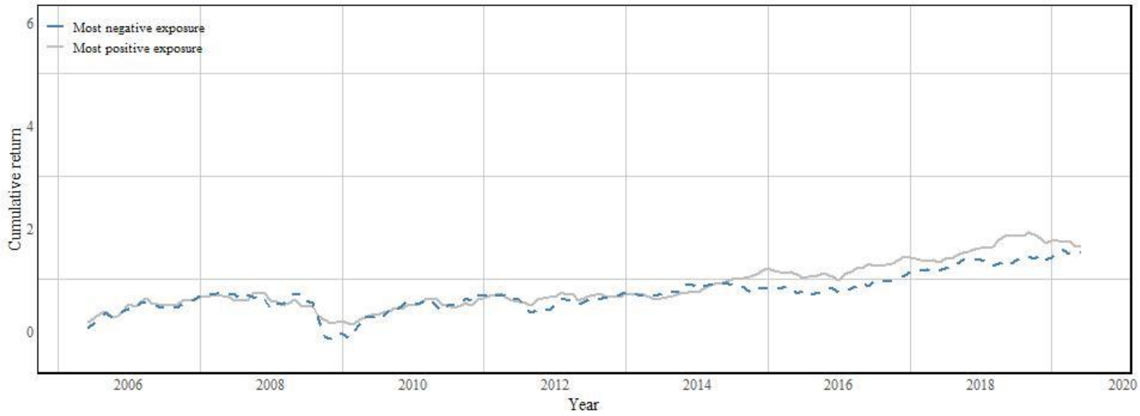
Cumulative excess returns of the two extreme EPU portfolios in the period of 2003-2019. Portfolios are simple sorted by exposure to EPU measured by online newspaper articles. The portfolios include the companies with the 1/3 most negative EPU exposure and the 1/3 most positive exposure. Portfolios are constructed at the end of June each year with weights held until next June. The associated t-test determines if there is a statistically significant difference in mean of returns between the two groups, with the null hypothesis stating that no such difference exists. A low p-value indicates evidence to reject the null hypothesis.



Mean most negative	Mean most positive	t-statistic	p-value
0.012	0.012	<-0.01	0.99

Figure 5.3: EPU Portfolio Spreads – Google Searches

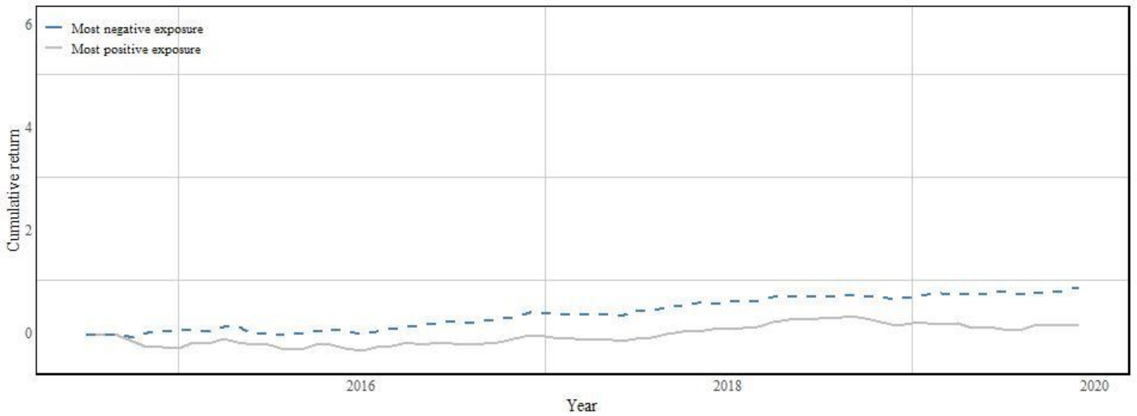
Cumulative excess returns of the two extreme EPU portfolios in the period of 2005-2019. Portfolios are simple sorted by exposure to EPU measured by Google searches. The portfolios include the companies with the 1/3 most negative EPU exposure and the 1/3 most positive exposure. Portfolios are constructed at the end of June each year with weights held until next June. The associated t-test determines if there is a statistically significant difference in mean of returns between the two groups, with the null hypothesis stating that no such difference exists. A low p-value indicates evidence to reject the null hypothesis.



Mean most negative	Mean most positive	t-statistic	p-value
0.009	0.010	-0.10	0.92

Figure 5.4: EPU Portfolio Spreads – Firm-specific

Cumulative excess returns of the two extreme EPU portfolios in the period of 2014-2019. Portfolios are simple sorted by firm-specific exposure to EPU. The portfolios include the companies with the 1/3 most negative EPU exposure and the 1/3 most positive exposure. Portfolios are constructed at the end of June each year with weights held until next June. The associated t-test determines if there is a statistically significant difference in mean of returns between the two groups, with the null hypothesis stating that no such difference exists. A low p-value indicates evidence to reject the null hypothesis.



Mean most negative	Mean most positive	t-statistic	p-value
0.013	0.002	1.32	0.19

5.2 Factor Realizations

To perform a factor analysis using the Fama-Macbeth framework, we need realization estimates of each factor at every point in time. These estimates are calculated as the value-weighted returns of the mimicking portfolios explained in Section 3.2.3. Summary statistics of the realization estimates are illustrated in Table 5.1. Each factor receives four different realization estimates due to the EPU measures covering differing time periods. When looking at the EPU factor realizations, we find a positive value when using our print newspaper-based measure. This is in line with the theoretical foundation, considering that the portfolio consists of buying stocks with the lowest exposure to EPU (most negative beta) and selling the stocks with the highest exposure to EPU (most positive beta). On the contrary, when using the online newspaper-based measure and the Google search-based measure of EPU, we obtain negative values for the mimicking portfolios. Nevertheless, when using aggregate measures, none of our mimicking portfolios proxying for risk related to policy uncertainty obtain significant values. Thus, we do not have statistical evidence to infer they are different from zero. However, the mimicking portfolio proxying for policy uncertainty when using our firm-specific measure of EPU does obtain statistically significant returns. Furthermore, the value is positive as expected from theory, stating that the firms with the most negative exposure to EPU obtain on average 1.4% higher returns per month than the companies with the highest positive exposure to policy uncertainty. Most of the factors used as controls have values in line with what we should expect, except from the operating profitability factor which is negative for all time periods. The factor proxying for risk related to size obtains values statistically significant from zero in all time periods, while the market factor appears significant for all time periods except the shortest one (2014-2019). Apart from these, only the profitability factor realization at the horizon of our online index (2003-2019) is statistically significant.

We find the realization estimates of the control factors generally to be in line with previous research conducted on the OSE. Like Næs et al. (2009), we find both MKT and SMB to be positive and statistically significant. However, while we find the value-weighted excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks to be significant for all time periods, Næs et al. (2009) only finds this effect to be significant from 1980-2000 (the study only investigate returns between 1980 and 2006). Ødegaard (2020b), however, finds a significantly positive effect for the period of 1980-2017. For our HML factor, we do not find that the excess returns are significantly different from zero. Contrary, Næs et al. (2009) does find a significantly positive value for the realization of the HML factor, however, only significant in the decade 1980-2000. Our findings are in line with Ødegaard (2020b) who finds the HML factor to be insignificant for the time periods 1980-2017 and 2000-2017. As we investigate returns on a different time horizon, the results cannot be compared directly. Regarding the realizations of the profitability factor, we only obtain a significant estimate at the horizon of 2003-2019. In contrast to theory, we find a negative estimate for a value-weighted portfolio of the most profitable firms minus a value-weighted portfolio of the least profitable firms. It does not exist published research regarding the realization of this factor using returns from the OSE. However, some master theses' finds the factor to be insignificant (Hoel & Mix, 2016; Bakken, 2019). Again, these studies are based on different time periods than the estimate coinciding with the horizon of our online article EPU measure and may therefore not be compared directly. Furthermore, we do not obtain significant estimates for the other time periods investigated. Regarding the CMA factor realization, we consistently obtain slightly positive, yet insignificant, estimates for all time periods. No published research exists for the investment factor and the research contributed from master theses is ambiguous (Hoel & Mix, 2016; Bakken, 2019). However, the thesis using the Compustat Global database does not find a significant realization of a value-weighted portfolio of conservative stocks minus a value-weighted portfolio of aggressive stocks (Hoel & Mix, 2016), which is in line with our findings.

Table 5.1: Factor Realizations

Average monthly excess returns for all factors in respective time periods. The t-statistics and associated p-values are obtained from testing whether the factor realizations are significantly different from zero. Stocks are allocated to mimicking portfolios at the end of June each year and the positions are held constant until the same period in the following year. The portfolios are constructed as stated in Section 3.2.3. MKT is the return of the market portfolio in excess of the risk-free rate. SMB is the average monthly excess return of a portfolio of small stocks minus a portfolio of large stocks, HML is the average monthly excess return of a value portfolio minus a growth portfolio, RMW is the average monthly excess return of a robust profitability portfolio minus a weak profitability portfolio, CMA is the average monthly excess return of a conservative investment portfolio minus an aggressive investment portfolio and EPU is the average monthly excess return of a negative EPU exposure portfolio minus a positive exposure EPU portfolio. Coefficients are noted as percentages.

	MKT	SMB	HML	RMW	CMA	EPU
Print articles (1992-2019)						
Mean	1.16	1.20	0.41	-0.63	0.29	0.43
t-statistic	3.60	4.84	1.26	-1.83	1.00	1.10
p-value	<0.01	<0.01	0.21	0.07	0.32	0.27
Online articles (2003-2019)						
Mean	1.22	0.98	0.38	-0.97	0.39	-0.35
t-statistic	3.34	3.09	0.94	-2.61	1.06	-0.79
p-value	<0.01	<0.01	0.35	<0.01	0.29	0.43
Google searches (2005-2019)						
Mean	0.94	0.97	0.27	-0.64	0.31	-0.16
t-statistic	2.39	2.83	0.63	-1.65	0.79	-0.29
p-value	0.02	<0.01	0.53	0.10	0.43	0.77
Firm-specific EPU: Textual analysis (2014-2019)						
Mean	0.75	0.89	-0.43	-0.51	1.00	1.40
t-statistic	1.86	2.01	-0.57	-0.79	1.42	2.15
p-value	0.07	0.05	0.57	0.43	0.16	0.04

5.3 Factor Exposures

To estimate the exposures of the test portfolios to various factors, we apply time series regressions in line with the Fama-Macbeth framework. Investigating if our test portfolios have statistically significant exposures to EPU will give a first impression of a potential linear relationship between policy uncertainty and expected returns. Since we have different test portfolios based on the various sorting mechanisms, the test portfolios will obtain one exposure per combination of sorting method, control model and measure of EPU. This implies 36 different specifications when illustrating estimates of factor exposures. As the findings regarding the control factors are not of main interest in this paper, and considering that they support what is established in earlier research, we only include the exposures related to different measures of EPU in the main analysis. We refer to the appendix for the tables including regression coefficients for all variables. The regression tables are structured to deal with one measure of policy uncertainty at a time. The time series regressions based on print newspaper articles, online newspaper articles, Google searches and the firm-specific measure of EPU will be discussed separately before summarizing the findings.

5.3.1 EPU Exposure: Print Articles

When measuring EPU by print newspaper articles, very few test portfolios obtain statistically significant estimated exposures to policy uncertainty. In Table 5.2 we observe that only one test portfolio at the Size-OP sort obtains a statistically significant exposure to policy uncertainty when controlling for the CAPM. When controlling for the Fama-French Three-factor model, two additional test portfolios obtain statistically significant estimates. The results when controlling for the Fama-French Five-factor model are mainly in line with the findings when controlling for the Three-factor model, with one less portfolio obtaining a significant estimate. Consequently, this implies that policy uncertainty is a risk factor that poorly explain variation in Norwegian stock returns when measuring EPU by print newspaper articles.

5.3.2 EPU Exposure: Online Articles

The time series regression estimates when using online newspaper articles as a measure of policy uncertainty are illustrated in Table 5.3. When controlling for the CAPM, three test portfolios obtain statistically significant exposures at the 5% confidence level. However, the majority of test portfolios attain insignificant exposures. The results do not improve much when adding the Three-factor model as control, with only one more test portfolio obtaining statistically significant exposure. The outcome when controlling for the Fama-French Five-factor does not change at all compared to the results controlled for the Three-factor. For all control model specifications, we obtain significant exposures to EPU at the Size-B/M and Size/OP sorting only. The result suggests that the relationship between expected returns and EPU measured by online newspaper articles is non-existent.

5.3.3 EPU Exposure: Google Searches

The exposures obtained when using our Google search-based measure of policy uncertainty is illustrated in Table 5.4. When controlling for the CAPM, three test portfolios obtain statistically significant exposures to EPU at the 5% confidence level. The results do not change much when including the Fama-French Three-factor as control, with only one more test portfolio obtaining a statistically significant value. We see that the situation is more or less the same when using the Fama-French Five-factor as controls. Furthermore, some of the significant exposures are associated with portfolios which fail to fulfil Ødegaard's (2020c) criteria of being considered well-diversified portfolios. In particular, the portfolio consisting of large-cap stocks with weak operating profitability is the weakest portfolio in regard to this objective, constituting an average of only four stocks throughout the time period. This test portfolio obtains significant estimated exposure when controlling for the Fama-French Three-and Five-factor models. The findings imply little statistical evidence for claiming this EPU measure to be a suitable variable for explaining variation in Norwegian stock returns.

5.3.4 EPU Exposure: Firm-specific Measure

When inspecting the test portfolios' exposure to the firm-specific EPU factor in Table 5.5, we see that very few obtain statistically significant exposures. The only ones with significant exposures, when controlling for the CAPM, are the portfolios with small-cap low book-to-market and small-cap aggressive investment stocks. Both portfolios experience issues with containing satisfying amounts of stocks throughout the time period. When adding the Fama-French Three-factor as a control, also the medium size neutral book-to-market portfolio obtain significant exposure. However, the other 24 out of 27 portfolios do not acquire significant values. The results are similar to the results controlled for the CAPM when controlling for the Fama-French Five-factor; only the small-cap low book-to-market portfolio and the small-cap aggressive investment portfolio obtain significant exposures. As stated, two of the portfolios which obtain significant estimates consists of few stocks. The portfolio of small-cap stocks with low book-to-market consists of only six stocks on average while the portfolio of small-cap stocks with aggressive investment consists of seven. Consequently, these are not regarded as well-diversified portfolios and consequently we should question the associated results. The result implies that the firm-specific measure of EPU does not explain returns in the Norwegian stock market.

5.3.5 EPU Exposure: Summary

From the time series regressions, none of our measures of policy uncertainty seem to explain Norwegian stock returns in a convincing manner. This is based on the test portfolios rarely obtaining statistically significant loadings on the EPU factor. The findings imply that there is no explicit linear relationship between expected returns and exposure to policy uncertainty. Ultimately, this implies that policy uncertainty does not seem to be a systematic risk factor for explaining returns at the OSE. Nonetheless, we continue our analysis by inspecting the different risk premia obtained from our cross-sectional regressions.

In Table A.4 to A.15, all estimates from the time series regressions are presented. At large, only the size and the market factors obtain exposures statistically significant for most test portfolios. Regarding our remaining control factors, these rarely obtain exposures with statistically significant values. As expected, and in line with previous research, we find that the test portfolios with the highest book-to-market characteristics obtain higher exposures to the HML factor than the test portfolios with low book-to-market characteristics. However, the obtained values are rarely statistically significant. Furthermore, we obtain statistically insignificant values for our model intercepts when using the Three-and Five-factor models, implying that the models capture most of the variation in returns. This is, however, not the case for the CAPM, stating that the CAPM combined with the EPU factor performs poorly when explaining returns in the Norwegian stock market. However, as the control factor estimates are in line with findings from previous studies on Norwegian stock returns, we do not dwell further into these results.

Table 5.2: Test Portfolios' Factor Exposures to EPU - Print Articles

EPU exposures obtained from the Fama-Macbeth time series regressions. The Five-factor model including the EPU factor for the time series regression is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between December 1992 and June 2019.

Size ↓	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Small	0.10	0.05	-0.02	0.10	0.04	-0.03	0.09	0.01	-0.04
2	-0.06	-0.01	-0.03	-0.08**	-0.02	-0.04	-0.07	-0.01	-0.05
Big	0.03	0.01	0.00	0.03	0.01	0.00	0.05	0.01	0.01
OP →	L	2	H	L	2	H	L	2	H
Small	0.09	0.03	0.05	0.08	0.02	0.04	0.08	0.01	0.01
2	-0.11**	-0.01	-0.02	-0.12***	-0.02	-0.04	-0.12***	-0.02	-0.04
Big	-0.09	0.02	0.02	-0.10	0.02	0.02	-0.12	0.03	0.03
INV →	L	2	H	L	2	H	L	2	H
Small	0.05	0.05	0.07	0.03	0.04	0.07	0.01	0.03	0.05
2	-0.10*	0.00	-0.04	-0.12***	-0.01	-0.05	-0.12***	0.00	-0.06
Big	0.01	0.03	0.00	0.01	0.03	0.00	0.01	0.04	-0.01

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table 5.3: Test Portfolios' Factor Exposures to EPU - Online Articles

EPU exposures obtained from the Fama-Macbeth time series regressions. The Five-factor model including the EPU factor for the time series regression is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between February 2003 and June 2019.

Size ↓	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Small	0.24**	0.14**	-0.01	0.21**	0.14**	-0.04	0.24**	0.13**	-0.04
2	-0.11*	-0.03	0.01	-0.12**	-0.02	-0.02	-0.11**	-0.02	-0.04
Big	0.07	-0.01	0.07	0.06	-0.01	0.05	0.06	-0.01	0.05
OP →	L	2	H	L	2	H	L	2	H
Small	0.25***	0.03	0.05	0.23***	0.02	0.03	0.22***	0.04	0.00
2	-0.06	-0.01	-0.05	-0.07	-0.03	-0.05	-0.06	-0.04	-0.06
Big	-0.09	0.01	-0.01	-0.10	-0.01	0.00	-0.11	-0.01	0.00
INV →	L	2	H	L	2	H	L	2	H
Small	0.12	0.06	0.19*	0.12	0.03	0.16	0.09	0.03	0.20*
2	-0.05	-0.02	-0.04	-0.06	-0.03	-0.05	-0.08	-0.02	-0.06
Big	0.03	0.02	0.06	0.02	0.02	0.07	0.02	0.03	0.05

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table 5.4: Test Portfolios' Factor Exposures to EPU - Google Searches

EPU exposures obtained from the Fama-Macbeth time series regressions. The Five-factor model including the EPU factor for the time series regression is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between June 2005 and June 2019.

Size ↓	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Small	0.11	0.10*	0.06	0.19	0.05	0.05	0.20*	0.07	0.06
2	0.04	0.03	-0.08	-0.04	-0.05	-0.07	-0.04	-0.04	-0.05
Big	-0.07	0.03*	-0.08	-0.08	0.04**	-0.07	-0.08	0.04**	-0.08
OP →	L	2	H	L	2	H	L	2	H
Small	0.21**	-0.01	-0.03	0.23***	-0.04	-0.02	0.24***	-0.05	0.00
2	0.03	-0.05	0.04	-0.04	-0.07*	-0.02	-0.04	-0.07	-0.02
Big	-0.11*	-0.08**	0.02	-0.14**	-0.07*	0.03	-0.14**	-0.07*	0.03
INV →	L	2	H	L	2	H	L	2	H
Small	0.12	-0.07	0.20*	0.09	-0.08	0.26**	0.11	-0.08	0.27**
2	0.04	0.00	-0.04	-0.03	-0.03	-0.10*	-0.02	-0.02	-0.10*
Big	-0.10**	-0.02	0.09*	-0.08	-0.03	0.06	-0.08*	-0.03	0.05

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table 5.5: Test Portfolios' Factor Exposures to Firm-specific EPU

EPU exposures obtained from the Fama-Macbeth time series regressions. The Five-factor model including the EPU factor for the time series regression is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t*. MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between July 2014 and June 2019.

Size ↓	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Small	-0.80**	0.06	0.07	-0.76**	0.02	0.03	-0.78**	0.03	-0.01
2	-0.02	-0.13	0.24*	-0.10	-0.21**	0.15	-0.03	-0.15	0.18
Big	-0.09	0.04	0.01	-0.12	0.06*	0.06	-0.15*	0.08*	-0.02
OP →	L	2	H	L	2	H	L	2	H
Small	-0.33	-0.09	0.14	-0.35	-0.10	0.08	-0.36	-0.11	0.10
2	0.10	-0.03	-0.07	0.00	-0.09	-0.17*	0.13	-0.06	-0.17
Big	0.00	0.03	-0.03	0.02	0.01	0.00	-0.09	0.01	-0.02
INV →	L	2	H	L	2	H	L	2	H
Small	0.30*	-0.05	-0.90***	0.25	-0.07	-0.89***	0.26	-0.15	-0.88***
2	-0.04	0.13	-0.10	-0.16	0.05	-0.15	-0.13	0.08	-0.04
Big	-0.11	0.02	0.01	-0.11	0.03	0.01	-0.11	0.02	0.04

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

5.4 EPU as Systematic Risk Premia

This section presents the estimated risk premia associated with policy uncertainty. We obtain the estimates from the cross-sectional regressions of the Fama-Macbeth framework. The estimates related to EPU are presented in this section while the risk premia estimates of the control variables are illustrated in the Appendix. Tables 5.6 to 5.9 illustrate the estimated risk premium associated with EPU when using the different measures of policy uncertainty and including various control models. The p-values associated with the coefficients indicate if there is a significant linear relationship between exposure to policy uncertainty and realized returns.

5.4.1 EPU as Systematic Risk Premia: Print Articles

In Table 5.6, the regression results are visualized when using print newspaper articles as measure of EPU. When using CAPM as control, risk premia estimates fluctuate somewhat between the different sorts, with positive risk premia at the Size/BM and Size/OP sorts and a negative risk premium when using the test portfolios constructed on size and investment characteristics. However, none of the estimates are considered statistically significant, and thus we cannot claim they are different from zero. When expanding our analysis to the Fama-French Three-factor model, we obtain positive estimates of risk premia for all sorts, in contrast to theory. None of the estimates are statistically significant at the 5% level. However, the estimate at the Size-B/M sorting is significant at the 10% level with a t-statistic of 1.94. Finally, when controlling for the Fama-French Five-factor model, we obtain similar estimates to the ones controlled for the CAPM. The estimate based on the Size-INV sort turn negative while the others remain positive. Again, none of the estimates are statistically significant. Thus, when measuring EPU by print newspaper articles, there is no significant linear relationship between exposure to policy uncertainty and expected returns.

5.4.2 EPU as Systematic Risk Premia: Online Articles

Table 5.7 presents risk premia estimates when measuring EPU by online newspaper articles. Like the print newspaper-based measure, we obtain positive estimates for the Size-B/M and Size-OP sort and a negative estimate for the Size-INV sort when controlling for the CAPM. However, none of the estimates obtain statistically significant estimates at the 5% confidence level, although the estimate obtained at the Size-OP sort does obtain significant estimates at the 10% level. When expanding our analysis by including the Fama-French Three-factor model, the estimates are still not statistically significant. At last, we obtain estimates when controlling for the Fama-French Five-factor model. At this setting, all risk premia estimates obtain positive values in contrast to the theoretical foundation. Nonetheless, neither of the estimates derive statistically significant values. Consequently, there are no signs of a linear relationship between assets' exposures to EPU and expected returns when using online newspaper articles as measure of policy uncertainty.

5.4.3 EPU as Systematic Risk Premia: Google Searches

We investigate the estimated risk premia associated with policy uncertainty when measuring it by Google searches in Table 5.8. When controlling for the CAPM, we see a similar pattern to our previous regressions; the estimates are positive at the Size-B/M and Size-OP sort and negative at the Size-INV sort. The risk premia estimate at the Size-OP sort obtain a statistically significant estimate of 4.64%, implying that increased exposure to EPU by one unit should increase monthly returns by 4.64%. The estimate may be considered large, but this is not surprising in regressions when the test portfolios have relatively modest exposures to the factor in question and we are trying to estimate a linear relationship. In order to obtain an increased exposure of 1, one would have to perform quite extensive long and short transactions. We find the estimates to be more or less in line when including the size and value effects. However, none of the risk premia estimates are deemed statistically significant at this specification. At last, we observe the risk premia estimates of the Fama-French Five-factor model. At this specification, all risk premia estimates turn positive. Again, the risk premium estimate at the Size-OP sorting obtains a positive value considered statistically significant. Although this advocates a linear relationship between exposure to policy uncertainty and expected returns when measuring EPU by Google searches, the estimate is positive and consequently in contrast to theory. Although one should not discard a relationship from this reasoning, as empirical relationships between various variables have turned out be in contrast to theory in literature, we need to keep in mind that the Google search-based measure of EPU was the one believed to be most affected by noise. This is partly due to its structure of including very few search words, as elaborated upon in Section 3.1.3. Adding that the Fama-French Three-factor is regarded as the model with the best fit for explaining stock returns of the three control models in the Norwegian stock market, we should keep in mind that the risk premium does not turn out significant at this specification. Furthermore, the risk premium is only estimated to be significant at one particular sort, and we would like to obtain significant results at several specifications in order to assess the variable as a systematic risk factor in Norwegian stock returns. Finally, few of the test portfolios obtained significant exposures to the EPU factor in the time series regressions. Thus, it is reasonable to interpret the findings with caution.

5.4.4 EPU as Systematic Risk Premia: Firm-specific Measure

In Table 5.9, the regression results are presented when using our firm-specific measure of EPU. We find the estimates to be consistently negative, as expected from theory, when controlling for the CAPM. However, none are statistically significant at the 5% level, even though the estimate at the Size-B/M sorting is significant at the 10% level. When extending our analysis by including the Fama-French Three-factor model as control, we find the Size-B/M and Size-INV sort estimates to obtain negative values and the estimate retrieved from the Size-OP sort being positive. Again, none of the risk premia estimates are statistically significant. However, it is worth noting that the estimate at the Size-B/M sort obtains a p-value of 8%, stating that we would only obtain such an extreme or more extreme value in 8% of the incidents given that the null hypothesis of the coefficient being zero, holds. We obtain negative estimates for each sort when including the Fama-French Five-factor as control. Again, we do not have statistical evidence to infer that the values are significantly different from zero. One should add that the exploration of the firm-specific measure as systematic risk premia in Norway is based on a limited amount of data (60 months of risk premia estimates), and thus the findings may be altered when inspecting returns over a longer period of time.

5.4.5 EPU as Systematic Risk Premia: Summary

In summary, the evidence for economic policy uncertainty being a systematic risk premium in the Norwegian stock market is indisputably weak, if not non-existing. Regarding both our newspaper-based measures, none of our risk premia estimates associated with EPU are considered statistically significant at the 5% level. Thus, we cannot claim that it exists a linear relationship between the proposed macro variable and expected returns. When using our Google-based measure of EPU, we do obtain some evidence of a relationship. However, only two out of nine risk premia estimates turn out statistically significant and resultingly this must be regarded as weak evidence. Furthermore, the relationship contrasts with theory and evidence from other markets. Even though anomalies are found to behave distinctly in different regional stock markets, the contrary findings make the results questionable. When using the firm-specific measure of EPU, we obtain negative risk premia estimates in 8 out of 9 estimates.

However, we do not have statistical evidence to infer that they are different from zero, and thus no statistically robust relationship holds. To conclude, we do not find a negative linear relationship between policy uncertainty and expected returns in the cross-section of Norwegian stock returns. In fact, we do not find strong evidence of economic policy uncertainty carrying a systematic risk premium at all. Consequently, we disregard the hypothesis of policy uncertainty carrying a negative risk premium in the Norwegian stock market.

Several explanations for this result may be considered. Given Norway's stable governance, transparent institutions and relatively short distance between political parties regarding exercise of authority, it could be the case that policy uncertainty is a variable of limited importance in the Norwegian stock market. On the other hand, it could also be the case that a relationship exists between policy uncertainty and expected returns in Norway, but we fail to capture it. Given that the marginal investor determines stock prices, it could be the case that our indices do not reflect policy uncertainty perceived by the marginal investor at the OSE. However, since we include four distinct methods of capturing EPU in our analysis, the study should be relatively robust, as at least one measure should likely be able to capture perceived policy uncertainty. Moreover, the reason for not finding a systematic link between policy uncertainty and expected returns may arise from our method of analysis. For one, it could be the case that the relationship between policy uncertainty and expected returns is a nonlinear function. Second, there exists no perfect model for capturing all variation in returns, and if it did, this study would contribute very little to literature. Consequently, we do not know if the applied factor models used in the analysis are ideal controls when estimating the effect of exposure to policy uncertainty. Third, the application of the Fama-Macbeth framework is not entirely unproblematic, leading to measurement issues associated with the concept of errors-in-variables. Furthermore, when estimating firm exposure to policy uncertainty, we are not able to control for other macro variables which may impact returns and likely correlate with policy uncertainty. These variables could potentially act as confounders and therefore influence the estimated relationship between returns and policy uncertainty. At last, some of our portfolios consists of few stocks due to the relatively small scale of the Norwegian stock market in terms of listings. It could be the case that the failure to create fully diversified portfolios contributes to the lack of findings.

The complete tables from the cross-sectional regressions are illustrated in Tables A.16 to A.19. The control variable estimates from the cross-sectional regressions are in line with earlier research on stock returns at the OSE. The size factor obtains statistically significant positive values in the Three-factor model for every specification apart from one, which is in line with both theory and established research. The market factor does in many cases obtain significant risk premiums in the CAPM and EPU framework. However, the sign of the risk premium is contrary to theory; we perceive a negative risk premium for the market factor in every case it obtains statistically significant values at the 5% confidence level. The result may be supported by research finding the beta anomaly to be persistent at Oslo Stock Exchange; the phenomena of low beta stocks achieving high abnormal returns relative to high beta stocks (Baker, M. et al., 2013; Støle & Rojahn, 2019). The linear relationship between exposure to the market and expected returns disappears when we control for more factors. The book-to-market factor does not obtain statistically significant risk premia estimates, although it attains significant estimates at the 10% level for some specifications. This is in line with Næs et al. (2009), which only finds the book-to-market factor to be significant from 1980 until 1989 (before our period of investigation). Regarding the operating profitability factor, we find it to be insignificant in every sorting method and time period apart from one; the Size-INV sort at the 2003-2019 time horizon. In contrast to theory, the factor obtains a negative estimate, implying that higher operating profitability should lead to lower expected returns when holding exposure to market, size, book-to-market, investment and EPU (measured by online articles) constant. However, for every other time horizon, the estimated risk premia are insignificant. In general, our estimates coincide with Hoel and Mix (2016) and Bakken (2019), finding operating profitability to not be priced in the cross-section of stock returns at the OSE. At last, the CMA factor obtains insignificant risk premia estimates for all specifications. Again, this is in line with the findings of Hoel and Mix (2016) and Bakken (2019).

Table 5.6: Estimated EPU Risk Premia – Print Articles

EPU premia are obtained as the averages of λ^n retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{n}_i$ is the EPU estimated exposure of asset *i* retrieved from the time series regression and λ_t^n is the obtained risk premia associated with the EPU factor at time *t*. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 1992 and 2019. Coefficients noted as percentages.

	CAPM	Three-factor	Five-factor
SIZE-B/M portfolios			
Mean	4.26	5.95	4.25
t-statistic	1.52	1.94	1.28
p-value	0.13	0.05	0.20
SIZE-OP portfolios			
Mean	1.64	2.35	2.39
t-statistic	0.69	0.71	0.73
p-value	0.49	0.48	0.46
SIZE-INV portfolios			
Mean	-0.58	4.82	-6.78
t-statistic	-0.25	1.01	-1.29
p-value	0.80	0.31	0.20

Table 5.7: Estimated EPU Risk Premia – Online Articles

EPU premia are obtained as the averages of λ^n retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{n}_i$ is the EPU estimated exposure of asset *i* retrieved from the time series regression and λ_t^n is the obtained risk premia associated with the EPU factor at time *t*. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 2003 and 2019. Coefficients noted as percentages.

	CAPM	Three-factor	Five-factor
SIZE-B/M portfolios			
Mean	1.44	3.54	2.09
t-statistic	0.82	1.50	0.92
p-value	0.42	0.13	0.36
SIZE-OP portfolios			
Mean	3.37	1.64	1.91
t-statistic	1.93	0.91	1.05
p-value	0.06	0.37	0.30
SIZE-INV portfolios			
Mean	-1.57	-0.80	1.18
t-statistic	-0.69	-0.36	0.55
p-value	0.49	0.72	0.58

Table 5.8: Estimated EPU Risk Premia – Google Searches

EPU premia are obtained as the averages of λ^n retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{n}_i$ is the EPU estimated exposure of asset *i* retrieved from the time series regression and λ_t^n is the obtained risk premia associated with the EPU factor at time *t*. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. The risk premia are based on stock returns from the OSE between 2005 and 2019. Coefficients noted as percentages.

	CAPM	Three-factor	Five-factor
SIZE-B/M portfolios			
Mean	4.20	3.36	4.81
t-statistic	1.51	1.34	1.73
p-value	0.13	0.18	0.09
SIZE-OP portfolios			
Mean	4.64	2.43	4.21
t-statistic	2.50	1.38	2.37
p-value	<0.01	0.17	0.02
SIZE-INV portfolios			
Mean	-0.23	-0.40	0.96
t-statistic	-0.12	-0.23	0.28
p-value	0.90	0.82	0.78

Table 5.9: Estimated Firm-specific EPU Risk Premia

EPU premia are obtained as the averages of λ^n retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{n}_i$ is the EPU estimated exposure of asset *i* retrieved from the time series regression and λ_t^n is the obtained risk premia associated with the EPU factor at time *t*. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. The risk premia are based on stock returns from the OSE between 2014 and 2019. Coefficients noted as percentages.

	CAPM	Three-factor	Five-factor
SIZE-B/M portfolios			
Mean	-3.10	-3.04	-2.69
t-statistic	-1.76	-1.81	-1.54
p-value	0.08	0.08	0.13
SIZE-OP portfolios			
Mean	-2.96	0.93	-1.53
t-statistic	-1.21	0.32	-0.46
p-value	0.23	0.75	0.64
SIZE-INV portfolios			
Mean	-2.49	-1.69	-1.18
t-statistic	-1.61	-1.21	-0.78
p-value	0.11	0.23	0.44

6 Conclusion

The relation between macro variables and stock returns is receiving increased attention in international research. Economic policy uncertainty is one of the variables which is being studied more closely in recent years based on its likely impact on investment opportunity sets. In this thesis, we expand existing research by investigating the role of policy uncertainty in the cross-section of Norwegian stock returns. We address the hypothesis that exposure to policy uncertainty should carry a negative risk premium through the research question:

Is economic policy uncertainty priced in Norwegian stock returns?

We investigate this by estimating the linear relationship between expected stock returns and exposure to economic policy uncertainty using the Fama-Macbeth framework. Estimates are controlled for the CAPM, the Fama-French Three-factor model and the Fama-French Five-factor model and conducted by applying test portfolios sorted on size and book-to-market, size and operating profitability and size and investment characteristics. Given that there is no universally accepted method of capturing policy uncertainty, we expand existing literature by applying four distinct measures to our analysis. The measures include EPU captured by (i) newspaper print articles, (ii) newspaper online articles and (iii) Google searches as aggregate methods and (iv) textual analysis on company annual reports as a firm-specific method.

Our risk premia estimates provide little evidence in favour of the hypothesis when capturing EPU by aggregate measures. For all three measures, we find the estimates to be mainly insignificant. However, the Google-based measure does obtain significantly positive estimates when controlling for the CAPM and the Fama-French Five-factor model for the portfolios sorted on size and operating profitability. Nevertheless, there are elements that makes the findings questionable. First, we only find significant estimates for one sorting mechanism and for a limited time period (2005-2019). Second, of the models used as controls, the Fama-French Three-factor model is likely the most suited factor model for explaining stock returns in Norway. Insignificant findings at this specification questions the relationship. At last, our Google search-based measure of capturing policy uncertainty is believed to be the measure most affected by biases. Consequently, the findings should be interpreted with caution. When aiming to capture firm-specific EPU by applying textual analysis to annual reports, we obtain nearly exclusively negative risk premia estimates in line with theory. Nevertheless, the estimates are insignificant and thus do not bring evidence to our hypothesis. The risk premia estimation is conducted at a short time period of 60 months due to a limited number of annual reports and would therefore be interesting to investigate at a longer horizon. In summary, we do not have sufficient evidence to infer that policy uncertainty carries a systematic risk premium in Norwegian stock returns.

The findings imply that the marginal investor at the Oslo Stock Exchange (i) does not allocate attention to policy uncertainty or (ii) considers policy uncertainty to be a state variable unnecessary to hedge against. This presupposes that our measures are able to capture perceived economic policy uncertainty in a satisfying way. The second reason could be linked to stable and transparent governance carried out by Norwegian institutions as well as the distance in political practice between government and opposition usually being relatively minor. It could also be the case that the lack of significant estimates arises from failure of capturing economic policy uncertainty as truly perceived by the marginal investor at the OSE. Even though our indices are inspired by methods from published research, they may be subject to biases, and therefore may not represent policy uncertainty as perceived by the marginal investor. However, by applying four distinct methods of capturing EPU, the analysis should be relatively robust. At last, the econometric methods applied may also affect our results. First, the control variables applied to the analysis may not be the true factors present in the Norwegian stock market and including other variables could alter the findings. Second, when obtaining risk premia estimates, we use an estimation technique which implies rise of the errors-in-variables bias. Other methods could lead to different conclusions. At last, when estimating firm exposure to policy uncertainty, we would ideally control for other phenomena likely to affect returns and correlate with levels of EPU, such as general economic uncertainty. These variables could therefore act as confounders in our estimation of firm exposure to policy uncertainty. Introducing such variables could improve the robustness of the findings by isolating the policy uncertainty effect.

For further research, we believe investigating how to best capture investor attention could prove useful. In literature, studies apply various measures when exploring the relation between expected stock returns and macro variables that cannot be quantified directly, such as economic policy uncertainty. Consequently, it could be very useful to shed light on the properties of different methods. Additionally, further research could expand our study by increasing investigated time horizon, especially for the firm-specific measure or further develop it by introducing new sources of data, e.g. other types of company updates. At last, other estimation techniques and control variables could be applied.

References

Aharoni, G., Grundy, B. D., & Zeng, Q. (2012). Stock Returns and the Miller-Modigliani Valuation Formula: Revisiting the Fama-French Analysis. *SSRN Electronic Journal*, *61*(0), 1–24. https://doi.org/10.2139/ssrn.1800603

Al-Thaqeb, S. A., & Algharabali, B. G. (2019). Economic policy uncertainty: A literature review. *Journal of Economic Asymmetries*, *20*(July), e00133. https://doi.org/10.1016/j.jeca.2019.e00133

Azqueta-Gavaldon, A., Hirschbühl, D., Onorante, L., & Saiz, L. (2020). Economic policy uncertainty in the euro area: an unsupervised machine learning approach. *ECB Working Paper*, *2359*.

Baker, M., Bradley, B., & Taliaferro, R. (2013). The low-risk anomaly: A decomposition into micro and macro effects. *Financial Analysts Journal*, *70*(2), 43–58. https://doi.org/10.2469/faj.v70.n2.2

Baker, S. R., Bloom, N., & Davis, S. J. (2013). Measuring Economic Policy Uncertainty. In *Chicago Booth Research Paper*, *13*(2). https://doi.org/10.1093/qje/qjw024

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636. https://doi.org/10.1007/978-3-642-19712-3-48

Baker, S. R., Bloom, N., & Davis, S. J. (2020). Economic Policy Uncertainty Index. Retrieved from https://www.policyuncertainty.com/

Bakken, T. (2019). *The Fama-French Five-Factor Model and Norwegian Stock Returns* (Master thesis, NTNU, Trondheim, Norway). Retrieved from https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2634414

Bali, T. G., Brown, S. J., & Tang, Y. (2017). Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics*, *126*(3), 471–489. https://doi.org/10.1016/j.jfineco.2017.09.005

Banz, R. W. (1981). The Relationship Between Return and Market Value of Common Stocks. *Journal of Financial Economics*, *9*(1) 3-18. https://doi.org/10.1016/0304-405X(81)90018-0

Basu, S. (1977). Investment Performance of Common Stocks in Relation To Their Price-Earnings Ratios: a Test of the Efficient Market Hypothesis. *The Journal of Finance* *32*(3), 663–682. https://doi.org/10.1111/j.1540-6261.1977.tb01979.x

Bonaime, A., Gulen, H., & Ion, M. (2018). Does policy uncertainty affect mergers and acquisitions? *Journal of Financial Economics*, *129*(3), 531–558. https://doi.org/10.1016/j.jfineco.2018.05.007

Bontempi, M. E., Golinelli, R., & Squadrani, M. (2016). A New Index of Uncertainty Based on Internet Searches: A Friend or Foe of Other Indicators? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2746346

Brogaard, J., & Detzel, A. L. (2014). The Asset Pricing Implications of Government Economic Policy Uncertainty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2075375

Brown, S., Lajbcygier, P., & Li, B. (2007). Going negative: What to do with negative book equity stocks. *Journal of Portfolio Management*, *35*(1), 95–102. https://doi.org/10.3905/JPM.2008.35.1.95

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Castelnuovo, E., & Tran, T. D. (2017). Google It Up! A Google Trends-based Uncertainty index for the United States and Australia. *Economics Letters, 161*(October), 149–153. https://doi.org/10.1016/j.econlet.2017.09.032

Cochrane, J. H. (2000). *Asset pricing: (Revised Edition*). Princeton University Press.

Cochrane, J. H., and C. L. Culp (2003). Equilibrium Asset Prices and Discount Factors: Overview and Implications for Derivatives Valuation and Risk Management, in P. Field, *Modern Risk Management: A History*. London, U.K., Risk Books.

Colak, G., Durnev, A., & Qian, Y. (2017). Political Uncertainty and IPO Activity: Evidence from U.S. Gubernatorial Elections. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2281269

Da, Z., Engelberg, J. & Gao, P. (2011). In Search of Attention. *Journal of Finance*, *LXVI*(5), 1461–1499.

Davis, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *Journal of Finance*, *55*(1), 389–406. https://doi.org/10.1111/0022-1082.00209

DellaVigna, S., & Hermle, J. (2014). Does Conflict of Interest Lead To Biased Coverage? Evidence From Movie Reviews. *Nber Working Paper Series*. https://doi.org/10.3386/w20661

Donadelli, M. (2015). Google search-based metrics, policy-related uncertainty and macroeconomic conditions. *Applied Economics Letters*, *22*(10), 801–807. https://doi.org/10.1080/13504851.2014.978070

Dumas, B., & Solnik, B. (1995). The World Price of Foreign Exchange Risk. *American Finance Association*, *50*(2), 445–479. https://doi.org/10.1111/j.1540-6261.1995.tb04791.x

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 1–54. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance*, *50*(1), 131–155. https://doi.org/10.1111/j.1540-6261.1995.tb05169.x

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, *82*(3), 491–518. https://doi.org/10.1016/j.jfineco.2005.09.009

Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, *105*(3), 457–472. https://doi.org/10.1016/j.jfineco.2012.05.011

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, *123*(3), 441–463. https://doi.org/10.1016/j.jfineco.2016.11.004

Fama, E. F., & MacBeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Politcal Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Friedman, M. (1968). The Role of Monetary Policy. *The American Economic Review*, *58*(1), 95–110. https://doi.org/10.4324/9781315133607-5

Gailbraith, J. K. (1977). *The age of uncertainty. A history of economic ideas and their consequences.* Houghton Mifflin Harcourt.

Griffin, J. M. (2001). Are the Fama and French Factors Global or Country-Specific? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.262647

Gulen, H., & Ion, M. (2013). Policy uncertainty and corporate investment. *Review of Financial Studies*, *29*(3), 523–564. https://doi.org/10.1093/rfs/hhv050

Hansen, L. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, *50*(4), 1029–1054. Retrieved from https://econpapers.repec.org/RePEc:ecm:emetrp:v:50:y:1982:i:4:p:1029-54

Hassan, T. A., Hollander, S., Van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, *134*(4), 2135–2202. https://doi.org/10.1093/qje/qjz021

Haugen, R. A., & Baker, N. L. (1996). Commonality In The Determinants Of Expected Stock Returns. *Journal of Financial Economics*, *41*(3), 401–439. https://doi.org/10.1016/0304-405X(95)00868-F

Hoel, A. R., & Mix, F. R. (2016). *How suitable is the Fama-French five-factor model for describing German and Norwegian stock returns?* (Master thesis, NHH, Bergen, Norway). Retrieved from https://openaccess.nhh.no/nhh-xmlui/handle/11250/2407462

Hopkins, D. J., Kim, E., & Kim, S. (2017). Does newspaper coverage influence or reflect public perceptions of the economy? *Research and Politics*, *4*(4). https://doi.org/10.1177/2053168017737900

Jens, C. E. (2017). Political uncertainty and investment: Causal evidence from U.S. gubernatorial elections. *Journal of Financial Economics*, *124*(3), 563–579. https://doi.org/10.1016/j.jfineco.2016.01.034

Jensen, M. C., Black, F., & Scholes, M. S., (1972) The Capital Asset Pricing Model: Some Empirical Tests. Michael C. Jensen, Studies in the Theory of Capital Markets, Praeger Publishers Inc., 1972. Retrieved from https://ssrn.com/abstract=908569

Kim, G. (2020). Sensationalism in Online News. Retrieved from https://www.sfm.url.tw/php/Papers/CompletePaper/013-1600620446.pdf

Klein, R. W., & Bawa, V. S. (1977). The effect of limited information and estimation risk on optimal portfolio diversification. *Journal of Financial Economics*, *5*(1), 89–111. https://doi.org/10.1016/0304-405X(77)90031-9

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, *49*(5), 1541–1578. https://doi.org/10.1111/j.1540-6261.1994.tb04772.x

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. https://doi.org/10.2307/1924119

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica: Journal of the Econometric Society*, *41*(5), 867–887. https://doi.org/10.2307/1913811

Miller, M., & Modigliani, F. (1961) Dividend Policy, Growth and the Valuation of Shares. *Journal of Business*, 34, 411-433. http://dx.doi.org/10.1086/294442

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *34*(4), 768–783. https://doi.org/10.2307/1910098

Nichol, E., & Dowling, M. M. (2014). Profitability and Investment Factors for UK Asset Pricing Models. *SSRN Electronic Journal*, 1–7. https://doi.org/10.2139/ssrn.2511166

Nie, H., Ruan, R., & Shen, J. (2020). Firm-Level Economic Policy Uncertainty, Firms' Investment and Financial Assets. https://dx.doi.org/10.2139/ssrn.3714229

Novy-Marx, R. (2010). The other side of value: Good Growth and the Gross Profitability Premium. *NBER Working Paper* (15940). https://doi.org/10.4324/9781315158037-8

Næs, R., Skjeltorp, J., & Ødegaard, B. A. (2009). What factors affect the Oslo Stock Exchange? *Norges Bank*, *02*, 12. Retrieved from https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/39616/1/JESW15_003.pdf

Oslo Børs (2020a). Utenlandsk eierskap på høyeste nivå siden finanskrisen. Retrieved from https://www.oslobors.no/Oslo-Boers/Om-Oslo-Boers/Nyheter-fra-OsloBoers/Utenlandsk-eierskap-paa-hoeyeste-nivaa-siden-finanskrisen

Oslo Børs (2020b). Choosing your market. Retrieved from https://www.euronext.com/en/raise-capital/how-go-public/choosing-market

Pagan, A. (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors. *Economics Department of the University of Pennsylvania and Institute of Social and Economic Research, Osaka University*, *25*(1), 221–247. https://doi.org/10.2307/2648877

Pastor, L, & Veronesi, P. (2012). Uncertainty about Government Policy and Stock Prices. *Journal of Finance*, *67*(4), 1219–1264. https://doi.org/10.1111/j.1540-6261.2012.01746.x

Petersen, M. A. (2007). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, *22*(1), 435–480. https://doi.org/10.1093/rfs/hhn053

Prokop, D. (2014). Von Neumann–Morgenstern utility function. Retrieved October 12, 2020, from https://www.britannica.com/topic/von-Neumann-Morgenstern-utility-function

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, *9*(1), 19–46. https://doi.org/10.1016/0304-405X(81)90019-2

Riekeles, H. (2017). Hvor stor offentlig sektor tåler vi? *Civita*. Retrieved from https://www.civita.no/assets/2017/07/Hvor-stor-offentlig-sektor-t%C3%A5ler-vi_civita-notat_18_2017.pdf

Rodrik, D. (1991). Policy uncertainty and private investment in developing countries. *Journal of Development Economics*, *36*(2), 229–242. https://doi.org/10.1016/0304-3878(91)90034-S

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

Sharpe, W. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *XIX*, 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Skoulakis, G. (2008). Panel Data Inference in Finance: Least-Squares vs Fama-MacBeth. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.891408

Stulz, R. M. (1990). Managerial discretion and optimal financing policies. *Journal of Financial Economics*, *26*(1), 3–27. https://doi.org/10.1016/0304-405X(90)90011-N

Støle, K. F., & Rojahn, F. (2019). *Slow but Steady Wins the Race.* (Master thesis, NHH, Bergen, Norway). Retrieved from https://openaccess.nhh.no/nhh-xmlui/handle/11250/2645199

Sæbø, J. K. (2008). Norwegian Stock Market Anomalies. *Idunn*, *22*(1), 1-21 ER. http://www.idunn.no/beta/2008/01/norwegian_stock_market_anomalies

Treynor, J. (1961). Market Value, Time and Risk. https://dx.doi.org/10.2139/ssrn.2600356

Van Den Bosch, F. A. J., & De Man, A.-P. (1994). Government's Impact on the Business Environment and Strategic Management. *Journal of General Management*, *19*(3), 50–59. https://doi.org/10.1177/030630709401900304

Walkup, B. (2016). The impact of uncertainty on payout policy. *Managerial Finance*, *42*(11), 1054–1072. https://doi.org/10.1108/MF-09-2015-0237

Ødegaard, B. A. (2020a). Asset pricing data at OSE. Retrieved from http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html

Ødegaard, B. A. (2020b). Empirics of the Oslo Stock Exchange. Asset pricing results. *University of Stavanger*.

Ødegaard, B. A. (2020c). Empirics of the Oslo Stock Exchange. Basic, descriptive, results 1980-2019. *University of Stavanger*.

Appendix

Table A.1: Dictionary for Marking Newspapers Concerned with EPU

Dictionary used to retrieve newspaper articles concerned with policy uncertainty. Used both for the print article-based measure and the online article-based measure. The criteria applied is that every article needs to include one word related to (i) economy, (ii) governmental policy and (iii) uncertainty, in practice one word from each column.

Economic	Policy	Uncertainty
økonomi*	norges bank	usikker*
	sentralbank*	usikre
	regjering*	usikkerhet*
	departement*	uro*
	regulering*	
	minister*	
	direktiv*	
	storting*	

Note: * implies that all suffixes of the respective terms are included.

Table A.2: Textual Analysis Dictionary

Dictionary used to measure firm-specific policy uncertainty. Norwegian terms are applied on annual reports written in Norwegian while English terms are applied on annual reports written in English.

Norwegian terms	English terms
norges bank	central bank*
minister*	government*
sentralbank*	ministry
regjering*	ministries
departement*	regulation*
regulering*	directive*
minster*	parliament
direktiv*	
storting*	
myndighete*	

Note: * implies that all suffixes of the respective terms are included.

Table A.3: Companies per Test Portfolio

Average number of companies per test portfolio. The difference between number of companies per measure is due to different time horizons, with more companies listed on the OSE later in the period investigated.

	Print articles			Online articles			Google searches			Textual Analysis		
Size ↓	(1992-2019)			(2003-2019)			(2005-2019)			(2014-2019)		
B/M →	L	2	H	L	2	H	L	2	H	L	2	H
Small	7	9	9	8	11	11	8	11	12	6	10	15
2	10	12	11	12	15	13	13	15	13	14	15	12
Big	9	12	6	11	14	6	12	15	6	11	17	4
OP →	L	2	H	L	2	H	L	2	H	L	2	H
Small	12	8	5	14	10	6	15	10	6	16	10	4
2	11	13	10	12	17	11	12	17	11	12	17	12
Big	3	12	11	3	14	14	4	14	15	3	15	15
INV →	L	2	H	L	2	H	L	2	H	L	2	H
Small	10	8	7	12	10	8	13	10	8	13	11	7
2	9	13	12	11	14	14	12	14	15	12	12	16
Big	6	13	8	7	16	9	7	17	9	5	18	9

Table A.4: Test Portfolios' Factor Exposures Sorted on SIZE-B/M - Print Articles

Factor exposures obtained from the Fama-Macbeth time series regressions using nine value-weighted portfolios sorted on size and book-to-market characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between December 1992 and June 2019.

	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.01***	0.01***	0.01	0.01*	0.00	0.01	0.01*	0.00
2	0.01***	0.01***	0.01**	0.00	0.00*	0.00	0.00	0.00*	0.00
Big	-0.00	-0.00***	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		β			β			β	
Small	1.02***	0.70***	0.86***	1.06***	0.74***	0.91***	1.06***	0.74***	0.91***
2	0.92***	0.72***	0.94***	0.97***	0.75***	0.97***	0.97***	0.75***	0.97***
Big	1.01***	0.89***	0.92***	1.01***	0.88***	0.92***	1.01***	0.88***	0.92***
					s			s	
Small				0.57***	0.62***	0.78***	0.56***	0.61***	0.78***
2				0.67***	0.45***	0.47***	0.67***	0.45***	0.46***
Big				0.03	-0.16***	0.19***	0.04	-0.16***	0.19***
					h			h	
Small				0.09	-0.03	0.11*	0.09	-0.05	0.11
2				-0.09*	-0.09**	0.10**	-0.08*	-0.09**	0.09*
Big				0.04	0.01	0.13**	0.06	0.01	0.13**
								r	
Small							-0.02	-0.09	-0.02
2							0.04	0.02	-0.06
Big							0.07*	-0.00	0.01
								c	
Small							-0.11	-0.07	-0.02
2							0.02	-0.04	-0.06
Big							0.06	0.01	0.03
		n			n			n	
Small	0.10	0.05	-0.02	0.10	0.04	-0.03	0.09	0.01	-0.04
2	-0.06	-0.01	-0.03	-0.08**	-0.02	-0.04	-0.07	-0.01	-0.05
Big	0.03	0.01	0.00	0.03	0.01	0.00	0.05	0.01	0.01

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.5: Test Portfolios' Factor Exposures Sorted on SIZE-OP - Print Articles

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and profitability characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon\,,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between December 1992 and June 2019.

	CAPM			Three-factor			Five-factor		
OP →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.02***	0.01***	0.00	0.01**	0.00	0.00	0.01**	0.00
2	0.01***	0.01***	0.01***	0.00	0.00	0.00	0.00	0.00	0.00
Big	-0.01*	-0.01	-0.00*	-0.01**	-0.00	0.00	-0.01**	-0.00	0.00
		β			β			β	
Small	0.99***	0.74***	0.69***	1.04***	0.78***	0.72***	1.04***	0.78***	0.72***
2	1.00***	0.83***	0.71***	1.05***	0.86***	0.74***	1.05***	0.86***	0.74***
Big	1.37***	0.96***	0.89***	1.39***	0.96***	0.88***	1.39***	0.96***	0.88***
					s			s	
Small				0.80***	0.63***	0.45***	0.79***	0.62***	0.43***
2				0.68***	0.48***	0.40***	0.69***	0.47***	0.40***
Big				0.24**	0.06	-0.18***	0.23**	0.06	-0.18***
					h			h	
Small				0.09	0.05	0.10	0.09	0.04	0.07
2				-0.05	0.05	-0.08**	-0.05	0.05	-0.08*
Big				-0.08	0.07**	-0.05**	-0.09	0.08**	-0.04
								r	
Small							-0.02	-0.03	-0.14*
2							0.01	0.00	-0.00
Big							-0.06	0.04	0.04
								c	
Small							-0.06	-0.01	-0.20***
2							0.02	-0.04	-0.02
Big							-0.05	-0.01	0.03
		n			n			n	
Small	0.09	0.03	0.05	0.08	0.02	0.04	0.08	0.01	0.01
2	-0.11**	-0.01	-0.02	-0.12***	-0.02	-0.04	-0.12***	-0.02	-0.04
Big	-0.09	0.02	0.02	-0.10	0.02	0.02	-0.12	0.03	0.03

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.6: Test Portfolios' Factor Exposures Sorted on SIZE-INV - Print Articles

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and investment characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset i and NMP_t is the realization of the EPU factor at time t. MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between December 1992 and June 2019.

	CAPM			Three-factor			Five-factor		
INV →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.02***	0.00	0.01**	0.01***	-0.00	0.01**	0.01***	-0.00
2	0.01***	0.00*	0.01***	0.00	0.00	0.01**	-0.00	0.00	0.00**
Big	0.00	-0.00**	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		β			β			β	
Small	0.89***	0.84***	0.84***	0.94***	0.88***	0.88***	0.94***	0.88***	0.88***
2	0.91***	0.80***	0.88***	0.98***	0.82***	0.91***	0.97***	0.83***	0.90***
Big	0.89***	0.91***	1.08***	0.90***	0.91***	1.07***	0.90***	0.91***	1.07***
					s			s	
Small				0.75***	0.57***	0.62***	0.74***	0.57***	0.61***
2				0.92***	0.37***	0.40***	0.92***	0.36***	0.40***
Big				0.13**	-0.10**	-0.01	0.13**	-0.09**	-0.01
					h			h	
Small				0.01	0.10*	0.11	-0.02	0.09*	0.09
2				-0.07	0.00	-0.01	-0.07	0.01	-0.02
Big				-0.00	0.00	0.04	0.00	0.01	0.04
								r	
Small							-0.11	-0.04	-0.05
2							-0.00	0.02	-0.03
Big							0.01	0.03	-0.01
								c	
Small							-0.14*	-0.06	-0.05
2							0.08	-0.07*	-0.02
Big							0.02	0.05	0.05
		n			n			n	
Small	0.05	0.05	0.07	0.03	0.04	0.07	0.01	0.03	0.05
2	-0.10*	0.00	-0.04	-0.12***	-0.01	-0.05	-0.12***	0.00	-0.06
Big	0.01	0.03	-0.00	0.01	0.03	0.00	0.01	0.04	-0.01

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.7: Test Portfolios' Factor Exposures Sorted on SIZE-B/M - Online Articles

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and book-to-market characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between February 2003 and June 2019.

	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.02***	0.01*	0.01*	0.01***	-0.00	0.01*	0.01***	-0.00
2	0.01***	0.01***	0.01**	0.00	0.01**	0.00	0.00	0.01***	0.00
Big	-0.00	-0.00**	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		β			β			β	
Small	0.83***	0.70***	0.89***	0.92***	0.83***	1.05***	0.93***	0.84***	1.05***
2	0.88***	0.71***	0.98***	1.02***	0.80***	1.07***	1.02***	0.81***	1.07***
Big	1.03***	0.98***	0.88***	1.05***	0.95***	0.94***	1.04***	0.95***	0.94***
					s			s	
Small				0.51***	0.66***	0.86***	0.52***	0.64***	0.86***
2				0.73***	0.45***	0.50***	0.73***	0.44***	0.48***
Big				0.12*	-0.14***	0.33***	0.12*	-0.14***	0.33***
					h			h	
Small				0.10	-0.06	0.09	0.11	-0.08	0.09
2				-0.07	-0.11**	0.14**	-0.06	-0.11**	0.11*
Big				0.07	-0.02	0.09	0.07	-0.02	0.10
								r	
Small							0.07	-0.06	-0.03
2							0.02	-0.03	-0.14**
Big							0.03	-0.02	0.02
								c	
Small							-0.06	-0.14**	-0.07
2							0.00	-0.10*	-0.17**
Big							0.07	0.01	0.03
		n			n			n	
Small	0.24**	0.14**	-0.01	0.21**	0.14**	-0.04	0.24**	0.13**	-0.04
2	-0.11*	-0.03	0.01	-0.12**	-0.02	-0.02	-0.11**	-0.02	-0.04
Big	0.07	-0.01	0.07	0.06	-0.01	0.05	0.06	-0.01	0.05

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.8: Test Portfolios' Factor Exposures Sorted on SIZE-OP - Online Articles

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and profitability characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between February 2003 and June 2019.

	CAPM			Three-factor			Five-factor		
OP →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.01***	0.01**	0.01*	0.01*	0.01	0.01*	0.01*	0.01
2	0.02***	0.01***	0.01**	0.01*	0.00	0.00	0.01**	0.00	0.00
Big	-0.00	-0.00	-0.00**	-0.01	-0.00*	-0.00	-0.01	-0.00	-0.00
		β			β			β	
Small	0.94***	0.55***	0.70***	1.09***	0.66***	0.77***	1.09***	0.67***	0.79***
2	0.95***	0.83***	0.75***	1.10***	0.92***	0.83***	1.10***	0.93***	0.83***
Big	1.16***	1.08***	0.98***	1.21***	1.11***	0.95***	1.20***	1.11***	0.95***
					s			s	
Small				0.76***	0.62***	0.43***	0.75***	0.63***	0.40***
2				0.73***	0.52***	0.37***	0.72***	0.51***	0.36***
Big				0.29***	0.20***	-0.16***	0.28***	0.19***	-0.15***
					h			h	
Small				0.01	0.00	0.14	0.00	0.02	0.10
2				-0.06	0.05	-0.08	-0.06	0.04	-0.09
Big				0.04	0.06	-0.02	0.03	0.06	-0.02
								r	
Small							-0.06	0.10	-0.19*
2							0.01	-0.06	-0.06
Big							-0.04	-0.00	0.01
								c	
Small							-0.08	0.01	-0.28***
2							-0.04	-0.10**	-0.08
Big							0.01	-0.03	0.04*
		n			n			n	
Small	0.25***	0.03	0.05	0.23***	0.02	0.03	0.22***	0.04	0.00
2	-0.06	-0.01	-0.05	-0.07	-0.03	-0.05	-0.06	-0.04	-0.06
Big	-0.09	0.01	-0.01	-0.10	-0.01	0.00	-0.11	-0.01	-0.00

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.9: Test Portfolios' Factor Exposures Sorted on SIZE-INV - Online Articles

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and investment characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between February 2003 and June 2019.

	CAPM			Three-factor			Five-factor		
INV →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.02***	0.00	0.01**	0.01**	-0.00	0.01**	0.01**	-0.00
2	0.01**	0.00	0.02***	0.00	0.00	0.01***	0.00	0.00	0.01***
Big	0.00	-0.00**	-0.00	0.00	-0.00*	-0.00	0.00	-0.00*	-0.00
		β			β			β	
Small	0.84***	0.61***	0.78***	0.98***	0.71***	0.89***	0.98***	0.72***	0.90***
2	0.88***	0.80***	0.85***	1.08***	0.86***	0.94***	1.07***	0.88***	0.94***
Big	1.04***	1.00***	1.12***	1.07***	0.99***	1.12***	1.06***	1.00***	1.11***
					s			s	
Small				0.69***	0.59***	0.63***	0.66***	0.59***	0.64***
2				1.00***	0.31***	0.46***	0.99***	0.30***	0.45***
Big				0.20***	-0.02	-0.01	0.20***	-0.01	-0.01
					h			h	
Small				-0.08	0.16**	0.09	-0.11	0.15**	0.11
2				-0.06	0.01	-0.01	-0.07	-0.00	-0.02
Big				0.06	0.01	-0.10*	0.07	0.02	-0.11**
								r	
Small							-0.16	-0.03	0.12
2							-0.09	-0.03	-0.06
Big							0.04	0.05	-0.08
								c	
Small							-0.15*	-0.10	-0.05
2							-0.05	-0.15**	-0.06
Big							0.08	0.02	0.09
		n			n			n	
Small	0.12	0.06	0.19*	0.12	0.03	0.16	0.09	0.03	0.20*
2	-0.05	-0.02	-0.04	-0.06	-0.03	-0.05	-0.08	-0.02	-0.06
Big	0.03	0.02	0.06	0.02	0.02	0.07	0.02	0.03	0.05

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.10: Test Portfolios' Factor Exposures Sorted on SIZE-B/M – Google Searches

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and book-to-market characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between June 2005 and June 2019.

	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.01***	0.01	0.02**	0.01	-0.00	0.02**	0.01	-0.00
2	0.01***	0.01***	0.01*	0.00	0.01**	0.00	0.00	0.01**	0.00
Big	-0.00	-0.00**	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		β			β			β	
Small	0.78***	0.62***	0.83***	0.79***	0.75***	1.00***	0.80***	0.76***	1.00***
2	0.87***	0.66***	0.92***	1.06***	0.80***	0.99***	1.06***	0.81***	1.00***
Big	1.07***	0.98***	0.86***	1.08***	0.95***	0.91***	1.08***	0.94***	0.91***
					s			s	
Small				0.36**	0.54***	0.88***	0.35**	0.52***	0.86***
2				0.73***	0.46***	0.42***	0.74***	0.45***	0.39***
Big				0.07	-0.14***	0.32***	0.07	-0.14***	0.33***
					h			h	
Small				0.29**	-0.04	0.11	0.30*	-0.05	0.11
2				-0.12*	-0.15***	0.10	-0.10	-0.14**	0.08
Big				0.01	0.00	0.09	-0.01	0.00	0.11
								r	
Small							-0.02	-0.09	-0.05
2							0.07	0.03	-0.13*
Big							-0.09	-0.02	0.11
								c	
Small							-0.13	-0.13*	-0.08
2							0.03	-0.06	-0.16**
Big							0.03	0.01	0.03
		n			n			n	
Small	0.11	0.10*	0.06	0.19	0.05	0.05	0.20*	0.07	0.06
2	0.04	0.03	-0.08	-0.04	-0.05	-0.07	-0.04	-0.04	-0.05
Big	-0.07	0.03*	-0.08	-0.08	0.04**	-0.07	-0.08	0.04**	-0.08

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.11: Test Portfolios' Factor Exposures Sorted on SIZE-OP – Google Searches

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and profitability characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t*. MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between June 2005 and June 2019.

	CAPM			Three-factor			Five-factor		
OP →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02**	0.01***	0.01**	0.01	0.00	0.01	0.01	0.00	0.01
2	0.02***	0.01**	0.01*	0.01**	0.00	0.00	0.01**	0.00	0.00
Big	-0.01*	-0.00	-0.00**	-0.01**	-0.00	-0.00	-0.01**	-0.00	-0.00
		β			β			β	
Small	0.88***	0.48***	0.68***	0.98***	0.60***	0.76***	0.98***	0.60***	0.78***
2	0.93***	0.74***	0.68***	1.11***	0.84***	0.80***	1.11***	0.85***	0.81***
Big	1.01***	1.08***	0.98***	1.07***	1.10***	0.95***	1.07***	1.10***	0.95***
					s			s	
Small				0.62***	0.55***	0.48***	0.59***	0.55***	0.44***
2				0.68***	0.46***	0.42***	0.68***	0.45***	0.41***
Big				0.19*	0.17***	-0.16***	0.19*	0.16***	-0.15***
					h			h	
Small				0.17*	-0.01	0.11	0.14	0.00	0.09
2				-0.11	0.02	-0.11*	-0.09	0.02	-0.12*
Big				-0.04	0.05	-0.01	-0.04	0.04	-0.01
								r	
Small							-0.16	0.04	-0.15
2							0.10	-0.01	-0.03
Big							0.02	-0.05	0.02
								c	
Small							-0.11	0.00	-0.28***
2							-0.03	-0.08	-0.03
Big							0.01	-0.00	0.03
		n			n			n	
Small	0.21**	-0.01	-0.03	0.23***	-0.04	-0.02	0.24***	-0.05	0.00
2	0.03	-0.05	0.04	-0.04	-0.07*	-0.02	-0.04	-0.07	-0.02
Big	-0.11*	-0.08**	0.02	-0.14**	-0.07*	0.03	-0.14**	-0.07*	0.03

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.12: Test Portfolios' Factor Exposures Sorted on SIZE-INV – Google Searches

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and investment characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon ,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between June 2005 and June 2019.

	CAPM			Three-factor			Five-factor		
INV →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02***	0.01***	0.01	0.01**	0.01**	0.00	0.01**	0.01**	0.00
2	0.01**	0.00	0.02***	0.00	0.00	0.01***	0.00	0.00	0.01***
Big	0.00	-0.00*	-0.00	0.00	-0.00*	-0.00	0.00	-0.00*	-0.00
		β			β			β	
Small	0.69***	0.63***	0.79***	0.81***	0.73***	0.86***	0.82***	0.73***	0.87***
2	0.78***	0.77***	0.84***	1.01***	0.85***	0.97***	1.01***	0.86***	0.97***
Big	1.00***	1.02***	1.11***	1.00***	1.03***	1.13***	0.99***	1.02***	1.12***
					s			s	
Small				0.54***	0.51***	0.60***	0.51***	0.50***	0.59***
2				0.97***	0.28***	0.45***	0.95***	0.27***	0.45***
Big				0.09	0.01	-0.04	0.10	0.02	-0.03
					h			h	
Small				0.02	0.04	0.26*	-0.01	0.03	0.28*
2				-0.04	-0.04	-0.10	-0.06	-0.03	-0.09
Big				0.08	-0.00	-0.09	0.08	0.00	-0.10*
								r	
Small							-0.16*	-0.09	0.03
2							-0.12	0.02	0.04
Big							-0.00	0.02	-0.06
								c	
Small							-0.17*	-0.05	-0.12
2							-0.08	-0.11	-0.02
Big							0.06	0.05	0.05
		n			n			n	
Small	0.12	-0.07	0.20*	0.09	-0.08	0.26**	0.11	-0.08	0.27**
2	0.04	0.00	-0.04	-0.03	-0.03	-0.10*	-0.02	-0.02	-0.10*
Big	-0.10**	-0.02	0.09*	-0.08	-0.03	0.06	-0.08*	-0.03	0.05

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.13: Test Portfolios' Factor Exposures Sorted on SIZE-B/M – Firm-specific EPU

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and book-to-market characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between July 2014 and June 2019.

	CAPM			Three-factor			Five-factor		
B/M →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.05**	0.02**	0.01	0.04**	0.01*	0.00	0.05**	0.02**	0.00
2	0.01*	0.01**	0.00	0.00	0.01	-0.00	0.00	0.01	-0.00
Big	0.00	-0.00**	0.01	0.00	-0.00*	0.01	0.00	-0.00	0.00
		β			β			β	
Small	1.52***	0.36	0.55**	1.53**	0.57**	0.80***	1.47**	0.50**	0.77***
2	0.77***	0.47***	1.15***	1.04***	0.75***	1.51***	1.10***	0.75***	1.51***
Big	0.36***	1.32***	0.50*	0.44***	1.24***	0.44	0.44***	1.23***	0.47
					s			s	
Small				0.51	0.61***	0.79***	0.43	0.52**	0.77***
2				0.45**	0.54***	0.88***	0.50***	0.51***	0.86***
Big				0.04	-0.14**	0.23	0.06	-0.16**	0.31
					h			h	
Small				0.47*	0.15	0.23**	0.46	0.14	0.22**
2				-0.13	-0.07	0.08	-0.11	-0.06	0.08
Big				-0.13*	0.03	0.37**	-0.13*	0.03	0.37**
								r	
Small							-0.22	-0.19	-0.14
2							0.29**	0.07	0.04
Big							-0.04	-0.01	0.01
								c	
Small							-0.22	-0.25*	-0.06
2							0.16	-0.07	-0.04
Big							0.04	-0.04	0.22
		n			n			n	
Small	-0.80**	0.06	0.07	-0.76**	0.02	0.03	-0.78**	0.03	-0.01
2	-0.02	-0.13	0.24*	-0.10	-0.21**	0.15	-0.03	-0.15	0.18
Big	-0.09	0.04	0.01	-0.12	0.06*	0.06	-0.15*	0.08*	-0.02

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.14: Test Portfolios' Factor Exposures Sorted on SIZE-OP – Firm-specific EPU

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and profitability characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between July 2014 and June 2019.

	CAPM			Three-factor			Five-factor		
OP →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.02*	0.02***	0.02**	0.02	0.02***	0.02	0.02*	0.02***	0.02
2	0.00	0.02***	0.01	-0.00	0.01**	0.00	-0.00	0.01**	-0.00
Big	-0.01	0.01*	-0.00**	-0.01	0.00	-0.00	-0.01	0.00	-0.00
		β			β			β	
Small	1.28***	0.23	0.35	1.45***	0.34**	0.69**	1.35***	0.33*	0.68*
2	0.89***	0.69***	0.67***	1.19***	0.92***	1.06***	1.25***	0.92***	1.09***
Big	0.58**	0.86***	1.07***	0.50	0.92***	0.97***	0.48	0.95***	0.97***
					s			s	
Small				0.68*	0.43***	1.01***	0.56	0.43**	0.99***
2				0.48***	0.52***	0.82***	0.49***	0.51***	0.85***
Big				-0.15	0.10	-0.21***	-0.12	0.14	-0.21***
					h			h	
Small				0.30	0.19**	0.26	0.28	0.19**	0.25
2				-0.18*	0.02	-0.05	-0.16*	0.02	-0.04
Big				0.03	-0.04	0.01	0.02	-0.04	0.01
								r	
Small							-0.33	-0.02	-0.01
2							0.34***	0.04	0.10
Big							-0.21	0.10	-0.04
								c	
Small							-0.35	-0.01	-0.08
2							0.08	-0.03	0.09
Big							0.04	0.11	-0.01
		n			n			n	
Small	-0.33	-0.09	0.14	-0.35	-0.10	0.08	-0.36	-0.11	0.10
2	0.10	-0.03	-0.07	0.00	-0.09	-0.17*	0.13	-0.06	-0.17
Big	0.00	0.03	-0.03	0.02	0.01	0.00	-0.09	0.01	-0.02

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.15: Test Portfolios' Factor Exposures Sorted on SIZE-INV – Firm-specific EPU

Factor exposures obtained from the Fama-Macbeth time series regression using nine value-weighted portfolios sorted on size and investment characteristics. The Five-factor model including the EPU factor is noted

$$R_t^i - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + n_i NMP_t + \epsilon,$$

where n_i is the EPU exposure of asset *i* and NMP_t is the realization of the EPU factor at time *t.* MKT is the excess return of the market portfolio, SMB is the excess return of a portfolio of small-cap stocks minus a portfolio of large-cap stocks, HML is the excess return of a portfolio of high book-to-market stocks minus a portfolio of low book-to-market stocks, RMW is the excess return of a portfolio of robust profitability stocks minus a portfolio of weak profitability stocks and CMA is the excess return of a portfolio of conservative investment stocks minus a portfolio of aggressive investment stocks. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. Stocks are allocated to three size groups and three groups of the B/M, OP and INV factors in the month change of June/July each year, thus representing implementable trading strategies. Asterisks indicate if the exposures are statistically significant. Exposures are estimated using stock returns from the OSE between July 2014 and June 2019.

	CAPM			Three-factor			Five-factor		
INV →	L	2	H	L	2	H	L	2	H
Size ↓		α			α			α	
Small	0.01	0.02***	0.04**	0.01	0.02**	0.04**	0.01	0.02**	0.04**
2	0.01*	-0.00	0.01**	0.00	-0.01	0.01*	0.00	-0.01	0.01*
Big	0.01**	-0.00*	-0.01	0.01**	-0.00	-0.01	0.01*	-0.00	-0.01
		β			β			β	
Small	0.73***	0.45**	1.07**	0.97***	0.64***	1.17**	0.91***	0.63***	1.05**
2	0.73***	0.98***	0.69***	1.16***	1.30***	0.88***	1.19***	1.31***	0.91***
Big	0.78***	0.89***	1.23***	0.76***	0.85***	1.25***	0.79***	0.84***	1.30***
					s			s	
Small				0.66**	0.66***	0.64	0.58**	0.70***	0.48
2				0.85***	0.70***	0.32**	0.88***	0.69***	0.30*
Big				0.00	-0.08	0.03	0.04	-0.09	0.08
					h			h	
Small				0.12	0.24**	0.40	0.12	0.24**	0.38
2				-0.08	-0.01	-0.08	-0.08	-0.01	-0.07
Big				0.03	0.01	0.01	0.04	0.01	0.02
								r	
Small							-0.15	-0.11	-0.36
2							0.16	0.08	0.23**
Big							0.08	-0.04	0.22*
								c	
Small							-0.21	0.07	-0.44
2							0.11	-0.00	-0.03
Big							0.10	-0.03	0.16
		n			n			n	
Small	0.30*	-0.05	-0.90***	0.25	-0.07	-0.89***	0.26	-0.15	-0.88***
2	-0.04	0.13	-0.10	-0.16	0.05	-0.15	-0.13	0.08	-0.04
Big	-0.11	0.02	0.01	-0.11	0.03	0.01	-0.11	0.02	0.04

Note: Significant codes: p<0.01:***, p<0.05:**, p<0.1:*

Table A.16: Estimated Factor Risk Premia – EPU Measured by Print Articles

Factor premia are obtained as the averages of λ^i retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{b}_i$, $\hat{s}_i$, $\hat{h}_i$, $\hat{r}_i$, $\hat{c}_i$ and $\hat{n}_i$ are the estimated factor exposures to MKT, SMB, HML, RMW, CMA and NMP respectively for asset *i* retrieved from the time series regression and λ_t^f is the obtained risk premia associated with the factor *f* at time *t*. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 1992 and 2019. Coefficients noted as percentages.

Panel A: CAPM and Fama-French Three-factor as Controls

	CAPM		Three-factor			
	MKT	EPU	MKT	SMB	HML	EPU
SIZE-B/M portfolios						
Mean	-1.05	4.26	1.08	2.36	-1.39	5.95
t-statistic	-1.05	1.52	0.79	5.66	-0.78	1.94
p-value	0.30	0.13	0.43	<0.01	0.43	0.05
SIZE-OP portfolios						
Mean	-1.09	1.64	-0.70	2.24	-1.12	2.35
t-statistic	-1.20	0.69	-0.76	5.10	-0.51	0.71
p-value	0.23	0.49	0.45	<0.01	0.61	0.48
SIZE-INV portfolios						
Mean	-3.62	-0.58	1.05	2.09	-1.82	4.82
t-statistic	-2.82	-0.25	0.65	4.41	-0.39	1.01
p-value	<0.01	0.80	0.51	<0.01	0.70	0.31

Panel B: Fama-French Five-factor as Control

	Five-factor					
	MKT	SMB	HML	RMW	CMA	EPU
SIZE-B/M portfolios						
Mean	1.22	2.06	-1.32	-0.71	-2.83	4.25
t-statistic	0.83	3.72	-0.61	-0.16	-0.61	1.28
p-value	0.41	0.00	0.54	0.87	0.54	0.20
SIZE-OP portfolios						
Mean	-0.89	2.28	-0.17	-1.65	2.33	2.39
t-statistic	-0.95	4.50	-0.07	-0.28	0.49	0.73
p-value	0.34	0.00	0.95	0.78	0.63	0.46
SIZE-INV portfolios						
Mean	-0.93	0.57	2.41	-7.78	-3.75	-6.78
t-statistic	-0.34	0.82	0.54	-1.40	-1.01	-1.29
p-value	0.73	0.41	0.59	0.16	0.31	0.20

Table A.17: Estimated Factor Risk Premia – EPU Measured by Online Articles

Factor premia are obtained as the averages of λ^i retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{b}_i$, $\hat{s}_i$, $\hat{h}_i$, $\hat{r}_i$, $\hat{c}_i$ and $\hat{n}_i$ are the estimated factor exposures to MKT, SMB, HML, RMW, CMA and NMP respectively for asset *i* retrieved from the time series regression and λ_t^f is the obtained risk premia associated with the factor *f* at time *t*. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 2003 and 2019. Coefficients noted as percentages.

Panel A: CAPM and Fama-French Three-factor as Controls

	CAPM		Three-factor			
	MKT	EPU	MKT	SMB	HML	EPU
SIZE-B/M portfolios						
Mean	-3.61	1.44	0.40	1.85	-2.12	3.54
t-statistic	-3.23	0.82	0.16	4.05	-0.79	1.50
p-value	<0.01	0.42	0.88	<0.01	0.43	0.13
SIZE-OP portfolios						
Mean	-1.01	3.37	-0.07	2.11	-1.65	1.64
t-statistic	-1.23	1.93	-0.08	4.80	-0.77	0.91
p-value	0.22	0.06	0.94	<0.01	0.44	0.37
SIZE-INV portfolios						
Mean	-2.83	-1.57	-2.75	1.43	-4.31	-0.80
t-statistic	-3.08	-0.69	-1.72	2.88	-1.62	-0.36
p-value	<0.01	0.49	0.09	<0.01	0.11	0.72

Panel B: Fama-French Five-factor as Control

	Five-factor					
	MKT	SMB	HML	RMW	CMA	EPU
SIZE-B/M portfolios						
Mean	3.66	0.57	-3.29	8.55	-10.55	2.09
t-statistic	1.22	0.76	-1.18	1.11	-1.69	0.92
p-value	0.22	0.45	0.24	0.27	0.09	0.36
SIZE-OP portfolios						
Mean	0.85	1.51	-4.60	3.40	-5.64	1.91
t-statistic	0.68	2.12	-1.65	0.57	-0.90	1.05
p-value	0.50	0.04	0.10	0.57	0.37	0.30
SIZE-INV portfolios						
Mean	3.37	0.63	8.09	-10.31	-3.62	1.18
t-statistic	0.74	0.82	1.33	-2.19	-0.83	0.55
p-value	0.46	0.41	0.19	0.03	0.41	0.58

Table A.18: Estimated Factor Risk Premia – EPU Measured by Google Searches

Factor premia are obtained as the averages of λ^i retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{b}_i$, $\hat{s}_i$, $\hat{h}_i$, $\hat{r}_i$, $\hat{c}_i$ and $\hat{n}_i$ are the estimated factor exposures to MKT, SMB, HML, RMW, CMA and NMP respectively for asset *i* retrieved from the time series regression and λ_t^f is the obtained risk premia associated with the factor *f* at time *t*. The CAPM controls for MKT while the Three-factor controls for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 2005 and 2019. Coefficients noted as percentages.

Panel A: CAPM and Fama-French Three-factor as Controls

	CAPM		Three-factor			
	MKT	EPU	MKT	SMB	HML	EPU
SIZE-B/M portfolios						
Mean	-1.56	4.20	-1.21	1.21	-0.04	3.36
t-statistic	-1.25	1.51	-0.79	2.52	-0.03	1.34
p-value	0.21	0.13	0.43	<0.01	0.98	0.18
SIZE-OP portfolios						
Mean	-1.23	4.64	0.21	2.33	-0.55	2.43
t-statistic	-1.55	2.50	0.22	4.40	-0.31	1.38
p-value	0.12	<0.01	0.83	<0.01	0.75	0.17
SIZE-INV portfolios						
Mean	-2.94	-0.23	-1.62	1.37	-1.69	-0.40
t-statistic	-3.31	-0.12	-1.16	2.44	-0.79	-0.23
p-value	<0.01	0.90	0.25	0.02	0.43	0.82

Panel B: Fama-French Five-factor as Control

	Five-factor					
	MKT	SMB	HML	RMW	CMA	EPU
SIZE-B/M portfolios						
Mean	2.40	0.26	-1.95	5.72	-8.88	4.81
t-statistic	0.91	0.32	-1.10	1.44	-1.63	1.73
p-value	0.36	0.75	0.27	0.15	0.11	0.09
SIZE-OP portfolios						
Mean	0.47	1.58	-0.18	5.89	-4.97	4.21
t-statistic	0.48	2.71	-0.07	1.64	-1.65	2.37
p-value	0.63	0.01	0.94	0.10	0.10	0.02
SIZE-INV portfolios						
Mean	-2.88	1.66	-2.58	-0.07	2.64	0.96
t-statistic	-1.03	1.79	-1.00	-0.02	0.48	0.28
p-value	0.30	0.07	0.32	0.98	0.63	0.78

Table A.19: Estimated Factor Risk Premia – Firm-specific Measure of EPU

Factor premia are obtained as the averages of λ^i retrieved from the Fama-Macbeth cross-sectional regressions

$$R_t^i - R_t^f = \alpha_{n,t} + \lambda_t^b \hat{b}_i + \lambda_t^s \hat{s}_i + \lambda_t^h \hat{h}_i + \lambda_t^r \hat{r}_i + \lambda_t^c \hat{c}_i + \lambda_t^n \hat{n}_i + \epsilon_i,$$

where $\hat{b}_i$, $\hat{s}_i$, $\hat{h}_i$, $\hat{r}_i$, $\hat{c}_i$ and $\hat{n}_i$ are the estimated factor exposures to MKT, SMB, HML, RMW, CMA and NMP respectively for asset *i* retrieved from the time series regression and λ_t^f is the obtained risk premia associated with the factor *f* at time *t*. The CAPM controls for MKT while the Three-factor controsl for MKT, SMB and HML and the Five-factor is an extension of the Three-factor with RMW and CMA as additional controls. The t-statistics and associated p-values are retrieved from a test of whether the coefficients are different from zero, thus indicating if a risk premium is statistically significant. Risk premia are obtained for each of the three sorting methods: Size-B/M, Size-OP and Size-INV. Risk premia are based on stock returns from the OSE between 2014 and 2019. Coefficients noted as percentages.

Panel A: CAPM and Fama-French Three-factor as Controls

	CAPM		Three-factor			
	MKT	EPU	MKT	SMB	HML	EPU
SIZE-B/M portfolios						
Mean	0.43	-3.10	0.38	1.14	1.61	-3.04
t-statistic	0.58	-1.76	0.50	1.89	1.01	-1.81
p-value	0.56	0.08	0.62	0.06	0.32	0.08
SIZE-OP portfolios						
Mean	-0.72	-2.96	0.92	1.77	3.23	0.93
t-statistic	-0.93	-1.21	0.96	2.28	1.29	0.32
p-value	0.36	0.23	0.34	0.03	0.20	0.75
SIZE-INV portfolios						
Mean	-2.40	-2.49	-1.61	1.51	1.38	-1.69
t-statistic	-2.31	-1.61	-1.41	2.31	0.56	-1.21
p-value	0.02	0.11	0.16	0.02	0.58	0.23

Panel B: Fama-French Five-factor as Control

	Five-factor					
	MKT	SMB	HML	RMW	CMA	EPU
SIZE-B/M portfolios						
Mean	0.04	0.76	2.82	3.96	-4.49	-2.69
t-statistic	0.04	1.04	1.11	0.92	-1.32	-1.54
p-value	0.97	0.30	0.27	0.36	0.19	0.13
SIZE-OP portfolios						
Mean	0.85	0.36	8.49	5.95	-1.94	-1.53
t-statistic	0.76	0.38	1.76	1.47	-0.47	-0.46
p-value	0.45	0.70	0.08	0.15	0.64	0.64
SIZE-INV portfolios						
Mean	-2.04	1.53	3.10	3.86	-3.56	-1.18
t-statistic	-1.73	2.38	0.88	1.40	-1.51	-0.78
p-value	0.09	0.02	0.38	0.17	0.14	0.44

R Script

This section presents the R-code used in this study. The code includes every step included in our analysis except some repetitive tasks. In these cases, we only include one example.

```
####------------------------------------------------------------------####
####               Section 1 - Install and load packages               ####
####------------------------------------------------------------------####

### It is meant to run one section at a time. Read instructions at
### the top of each section for relevant information.

#load packages
library(data.table)
library(dplyr)
library(lubridate)
library(quantmod)
library(quanteda)
library(stringi)
library(readtext)
library(tidyverse)
library(readxl)
library(plyr)

####------------------------------------------------------------------####
####                      Section 2 - Format data                      ####
####------------------------------------------------------------------####

#read in data from CSV-files
stockdata_daily <- read.csv("compustat_stock.csv")
rf <- read.csv("Risk free.csv")
rf <- rf [-1,]
Index_data <- read.csv("Index.csv")

####--------------------------Risk free rate--------------------------####

# rename columns
rf <- rf %>%
  dplyr::rename(
    Date = Forward.looking.risk.free.rates,
    RiskFreeRate = X1m..estimated.from.govmt.securities.and.NIBOR.)

# change variable Date to date format
rf$Date <- as.Date(rf[["Date"]],"%Y%m%d")

# create year and month column from Date
rf <- rf %>%
  dplyr::mutate(Year = lubridate::year(Date),
                Month = lubridate::month(Date))

# change variable types to numeric
rf$RiskFreeRate <- as.numeric(levels(rf$RiskFreeRate)[rf$RiskFreeRate])

# subset dataframe
rf <- subset(rf, select = -Date)

####-------------------------Index return data------------------------####
```

```r
#rename variables
Index_data <- Index_data %>%
  dplyr::rename(
    Date = date)

#choose data to carry forward
Index_data <- subset(Index_data , select = c("Date","VW"))

# change variable Date to date format
Index_data <- transform(Index_data, Date = as.Date(as.character(Date),
                                                    "%Y%m%d"))


# create year and month column from Date
Index_data <- Index_data %>%
  dplyr::mutate(Year = lubridate::year(Date),
                Month = lubridate::month(Date))

# calculate market excess return
Index_data <- merge(Index_data, rf, by = c("Year", "Month"))
Index_data$VW_INDEX <- Index_data$VW - Index_data$RiskFreeRate
Index_data <- subset(Index_data, select = -VW)

####-------------------------------Stock data-------------------------####

# rename variables
stockdata_daily <- stockdata_daily %>%
  dplyr::rename(
    TradeDate = datadate,
    company = conm,
    Currency = curcdd,
    AdjustmentFactor = ajexdi,
    Price = prccd,
    ISIN = isin,
    SIC = sic,
    IssueType = tpci,
    SharesOutstanding = cshoc)

#change format of column from factor to date
stockdata_daily$TradeDate <-
  as.Date(as.character(stockdata_daily$TradeDate))

#extract year and month from "Date" to separate columns
stockdata_daily <- stockdata_daily %>%
  dplyr::mutate(Year = lubridate::year(TradeDate),
                Month = lubridate::month(TradeDate),
                Day = lubridate::day(TradeDate))

#create unique ID by combining gvkey and IID
stockdata_daily$ID <- as.factor(paste(stockdata_daily$gvkey,
                                      stockdata_daily$iid, sep = ""))

#keep only last observation each month for each stock
stockdata_monthly <- stockdata_daily %>%
  group_by(ID, Year, Month) %>%
  slice(which.max(Day))

# keep common stocks
stockdata_monthly <- subset(stockdata_monthly,
                            stockdata_monthly$IssueType == c(0))

#remove financial firms
```

```
stockdata_monthly <- stockdata_monthly[!(stockdata_monthly$SIC %in%
                                         6000:6999), ]

# keep only stocks at OSE
stockdata_monthly <- subset(stockdata_monthly,
                            stockdata_monthly$exchg == 228)

# only keep observations with trading day 25 or higher
stockdata_monthly <- subset(stockdata_monthly, stockdata_monthly$Day >= 25)

# subset dataframe
stockdata_monthly <- subset(stockdata_monthly, select = -c(cshtrd, exchg,
                                              IssueType, fyrc, SIC))

# adjust prices for splits and dividends
stockdata_monthly$PriceAdjusted <-
  stockdata_monthly$Price/stockdata_monthly$AdjustmentFactor

# make complete dataframe with observations each
# month each year for all stocks
dateseq <- as.data.frame(seq(as.Date("1986-01-01"),as.Date("2020-12-31"),
                             by = "month"))
dateseq <- dateseq %>%
  dplyr::mutate(Year = lubridate::year(`seq(as.Date("1986-01-01"),
                                  as.Date("2020-12-31"), by = "month")`),
                Month = lubridate::month(`seq(as.Date("1986-01-01"),
                                  as.Date("2020-12-31"), by = "month")`))
dateseq <- subset(dateseq, select = c("Year", "Month"))

Company_List <- unique(stockdata_monthly[, c("company", "ID")])
Company_date <- merge(dateseq, Company_List)
stockdata_monthly <- merge(stockdata_monthly, Company_date, all = TRUE)

# calculate monthly returns
stockdata_monthly <- data.table(stockdata_monthly)
stockdata_monthly[, MonthlyReturn := Delt(PriceAdjusted), by = ID]

# estimate MCAP
stockdata_monthly$MCAP <-
  stockdata_monthly$Price*stockdata_monthly$SharesOutstanding

# keep dataframe with penny stocks for later
Penny_Stock_file <- subset(stockdata_monthly, select = c("Price",
                                  "MonthlyReturn", "ID","Year", "Month"))

# remove penny stocks
Price_low <- subset(stockdata_monthly, Price < 1)
stockdata_monthly <- anti_join(stockdata_monthly,
                               Price_low, by = c("ID", "Year"))
stockdata_monthly <-
  stockdata_monthly[!is.na(stockdata_monthly$Price), ]

# remove all observations in a year were a stock has
# been valued below 1MNOK MCAP
MCAP_low <- subset(stockdata_monthly,stockdata_monthly$MCAP < 1000000)
stockdata_monthly <- anti_join(stockdata_monthly,
                               MCAP_low, by = c("ID", "Year"))

# subset dataframe
stockdata_monthly <- subset(stockdata_monthly, select = -c(Currency,
                      AdjustmentFactor,iid, TradeDate, SharesOutstanding,
```

```
                          Price, PriceAdjusted))

# Include risk free rate and calculate monthly excess return
stockdata_monthly <- merge(stockdata_monthly, rf, by = c("Year", "Month"),
                           all = T)
stockdata_monthly$ExcessReturn <-
  stockdata_monthly$MonthlyReturn - stockdata_monthly$RiskFreeRate

# save files
save(stockdata_monthly, file = "stockdata_monthly.Rdata")
save(Index_data, file = "Index_data.Rdata")
save(rf, file = "rf.Rdata")
save(Penny_Stock_file, file = "Penny_Stock_file.Rdata")



####------------------------------------------------------------------####
####                Section 3 - Format accounting data                ####
####------------------------------------------------------------------####

#read in data
load("stockdata_monthly.Rdata")
accounting_data  <- read.csv("account data.csv", sep = ",",
                             stringsAsFactors = FALSE)
exchange_USD <- read_csv2("USD_NOK.csv")
exchange_EUR <- read.csv2("EUR_NOK.csv")

####-----------------------Exchange rate data-------------------------####

# change format to date
exchange_EUR <- transform(exchange_EUR, TIME_PERIOD =
                            as.Date(as.yearmon(TIME_PERIOD)))
exchange_USD <- transform(exchange_USD, TIME_PERIOD =
                            as.Date(as.yearmon(TIME_PERIOD)))

# extract year and month to separate columns
exchange_EUR <- exchange_EUR %>%
  dplyr::mutate(Year = lubridate::year(TIME_PERIOD),
                Month = lubridate::month(TIME_PERIOD))

exchange_USD <- exchange_USD %>%
  dplyr::mutate(Year = lubridate::year(TIME_PERIOD),
                Month = lubridate::month(TIME_PERIOD))

# subset dataframe
exchange_EUR <- subset(exchange_EUR, select = c("OBS_VALUE", "Year"))
exchange_USD <- subset(exchange_USD, select = c("OBS_VALUE", "Year"))

#rename variables and create currency variable
exchange_EUR <- exchange_EUR %>%
  dplyr::rename(
    EUR = OBS_VALUE)

exchange_USD <- exchange_USD %>%
  dplyr::rename(
    USD = OBS_VALUE)

exchange_EUR$Currency <- "EUR"
exchange_USD$Currency <- "USD"

####---------------------------Accounting data------------------------####
```

```
#rename variables
accounting_data <- accounting_data %>%
  dplyr::rename(
    Currency = curcd,
    Year = fyear,
    Date = datadate,
    TotalAssets = at,
    COGS = cogs,
    TotalLiabilities = lt,
    TotalRevenue = revt,
    StockholdersEquity = seq,
    DeferredTaxes = txditc,
    TotalInterestExpense = xint,
    LongTermDebtExpense = xintd,
    OPEX = xopr,
    OtherOPEX = xopro,
    SGA = xsga,
    StockExchange = exchg,
    ISIN = isin,
    company = conm,
    CountryCode = fic,
    SIC = sic,
    Month = fyrc)

#change format of Date variable from factor to date
accounting_data$Date <- dmy(accounting_data$Date)

#remove financial firms
accounting_data <- accounting_data[!(accounting_data$SIC %in%
                                       6000:6999), ]

# keep only stocks at OSE
accounting_data <- subset(accounting_data,
                          accounting_data$StockExchange == 228)

# find and remove duplicate firms
n_occur <- data.frame(table(accounting_data$gvkey, accounting_data$Year))
n <- n_occur[n_occur$Freq > 1,]

# remove the duplicate that have the most NAs
accounting_data <- accounting_data[!(accounting_data$gvkey == "245498" &
                                accounting_data$Date ==  "2006-12-31"),]
accounting_data <- accounting_data[!(accounting_data$gvkey == "282118" &
                                accounting_data$Date ==  "2007-06-30"),]

# remove Roxar ASA as it is duplicated
accounting_data <- accounting_data[!(accounting_data$gvkey == "243374"),]

# merge accounting and currency data
accounting_data <- merge(accounting_data, exchange_EUR, by =
                          c("Year", "Currency"), all = TRUE)
accounting_data <- merge(accounting_data, exchange_USD, by =
                          c("Year", "Currency"), all = TRUE)

# remove Stepstone ASA 1998 since they operate in EUR before
# EUR started in january 1999
accounting_data <- accounting_data[!(accounting_data$gvkey == "235557" &
                                       accounting_data$Year ==  "1998"),]

#create a column with currency values
```

```
accounting_data <- accounting_data %>%
  mutate(CUR_VAL = case_when(
    Currency == "EUR" ~ accounting_data$EUR,
    Currency == "USD" ~ accounting_data$USD,
    Currency == "NOK" ~ 1
  ))

# convert all accounting values to NOK
accounting_data <- accounting_data %>%
  mutate(
    TotalAssets = TotalAssets*CUR_VAL,
    COGS = COGS * CUR_VAL,
    TotalLiabilities = TotalLiabilities * CUR_VAL,
    TotalRevenue = TotalRevenue * CUR_VAL,
    StockholdersEquity = StockholdersEquity*CUR_VAL,
    teq = teq*CUR_VAL,
    DeferredTaxes = DeferredTaxes * CUR_VAL,
    TotalInterestExpense = TotalInterestExpense * CUR_VAL,
    LongTermDebtExpense = LongTermDebtExpense * CUR_VAL,
    OPEX = OPEX * CUR_VAL,
    OtherOPEX = OtherOPEX *CUR_VAL,
    SGA = SGA * CUR_VAL)

# subset dataframe
accounting_data <- subset(accounting_data, select =
                  -c(Currency, indfmt, datafmt, consol, popsrc,
                     StockExchange, teq, costat, CountryCode, SIC,
                     EUR, USD, CUR_VAL, LongTermDebtExpense))


####-----------------------------------------------------------------####
####                Section 4 - Creating sorting variables           ####
####-----------------------------------------------------------------####

# omit companies with asset less or equal to zero
accounting_data <- data.table(accounting_data)
accounting_data <- subset(accounting_data, TotalAssets > 0)

#create book equity column
accounting_data$deftax <- replace_na(accounting_data$DeferredTaxes, 0)
accounting_data$BookEquity <- accounting_data$StockholdersEquity +
  accounting_data$DeferredTaxes
accounting_data$BookEquity <- replace_na(accounting_data$BookEquity, 0)
accounting_data$BookEquity = ifelse(accounting_data$BookEquity == 0,
          accounting_data$TotalAssets - accounting_data$TotalLiabilities
          + accounting_data$DeferredTaxes, accounting_data$BookEquity)

#remove observations with negative, zero or NA in book equity
accounting_data <- subset(accounting_data, accounting_data$BookEquity > 0)

#calculate profitability
accounting_data$OP <-
  ((accounting_data$TotalRevenue - accounting_data$OPEX-
     accounting_data$TotalInterestExpense)/ accounting_data$BookEquity)

# make complete time frame to make yearly investment calculation
accounting_data <- complete(accounting_data, company, Year)

# calculate investment variable
accounting_data <- data.table(accounting_data)
accounting_data <- accounting_data[order(company, Year)]
```

```
INVEST = function(x) (((x) - lag(x))/lag(x))
accounting_data[, INVEST := INVEST(TotalAssets), by = company]

# subset dataframe and omit NAs
accounting_data <- subset(accounting_data, select =
              c("Year", "gvkey", "company", "BookEquity", "OP", "INVEST"))
accounting_data <- na.omit(accounting_data)

# extract MCAP from stock data
df2 <- subset(stockdata_monthly, Month == 12, select =
                c("Year", "company", "MCAP", "ID"))
df4 <- subset(stockdata_monthly, Month == 6, select = c("Year", "MCAP", "ID"))

# merge MCAP for december with accounting data
sorting_variables <- merge(df2, accounting_data, by =
                             c("Year", "company"))

# calculate Book-to-market variable
sorting_variables$BM <- (sorting_variables$BookEquity /
                           sorting_variables$MCAP)

# add one year to lag sorting variables from t-1 to t
sorting_variables$Year <- (sorting_variables$Year +1)

sorting_variables <- subset(sorting_variables, select =
                              -c(MCAP, gvkey, company))

# make file with BE, OP, Investment and MCAP
df4 <- df4 %>%
  dplyr::rename(
    SIZE = MCAP)

sorting_variables <- merge(sorting_variables, df4, by =
                             c("Year", "ID"), all = T)
sorting_variables <- na.omit(sorting_variables)


####------------------------------------------------------------------####
####                 Section 5 - Forming test portfolios              ####
####------------------------------------------------------------------####

# cut in 0.3 and 0.7 quantiles to divide stocks into groups for each
# sorting variable
sorting_variables <- ddply(sorting_variables, .(Year), mutate,
                     SIZE_gr = cut(SIZE, quantile(SIZE, c(0, 0.3, 0.7, 1)),
                                   labels = c(1,2,3), include.lowest = T),
                     BM_gr = cut(BM, quantile(BM, c(0, 0.3, 0.7, 1)),
                                 labels = c(1,2,3), include.lowest = T),
                     OP_gr = cut(OP, quantile(OP, c(0, 0.3, 0.7, 1)),
                                 labels = c(1,2,3), include.lowest = T),
                     INVEST_gr = cut(INVEST, quantile(INVEST,
                                                      c(0, 0.3, 0.7, 1)),
                                 labels = c(1,2,3), include.lowest = T))

#assign stocks to test portfolios
sorting_variables <- transform(sorting_variables ,
                    SIZE_BM = paste0(SIZE_gr, BM_gr),
                    SIZE_INVEST = paste0(SIZE_gr, INVEST_gr),
                    SIZE_OP = paste0(SIZE_gr, OP_gr))

# expand each observation to compare yearly portfolio updates with
```

```
# monthly company returns
Month = expand.grid(Year = unique(sorting_variables$Year), Month = 1:12)
TestAssetData <- left_join(sorting_variables, Month, by = "Year")

# create a column to denote what monthly returns to sum from july to june
TestAssetData$return_Year = ifelse(TestAssetData$Month > 6,
                            TestAssetData$Year, TestAssetData$Year+1)

# rename column and merge dataframes
stockdata_monthly <- stockdata_monthly %>%
  dplyr::rename(
    return_Year = Year)

TestAssetData <- merge(TestAssetData, stockdata_monthly,
                       by = c("return_Year", "Month", "ID"))

#change back column name
stockdata_monthly <- stockdata_monthly %>%
  dplyr::rename(
    Year = return_Year)

#save dataset
TestAssetData <- as.data.frame(TestAssetData)
save(TestAssetData, file = "TestAssetData.Rdata")


####-----------------------------------------------------------------####
####                 Section 6 - Estimating EPU-betas                ####
####-----------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Print" or "Online"
### to run for other measures.

# load datasets
EPU_Index <- read_excel("EPU_Print.xlsx")
load("TestAssetData.Rdata")

# remove and rename columns
EPU_Index <- subset(EPU_Index, select = -c(Innovations, Index))
EPU_Index <- EPU_Index %>%
  dplyr::rename(
    EPU_Index = Percent)

# change format to date
EPU_Index$Date <- as.Date(EPU_Index[["Date"]],"%M%Y")

# extract year and month to separate columns
EPU_Index <- EPU_Index %>%
  dplyr::mutate(return_Year = lubridate::year(Date),
                Month = lubridate::month(Date))

TestAssetData <- merge(TestAssetData, EPU_Index, by =
                         c("return_Year", "Month"))

# create column to store EPU betas
TestAssetData$beta_monthly <- NA

# No. of observations in rolling regression
v <- 36

# No of observations of each stock requred in rolling regression
```

```
w <- 18

# set up dataframe
complete_Data <- TestAssetData
complete_Data <- complete(complete_Data, ID, Date)
complete_Data <- tibble::rowid_to_column(complete_Data, "RowNr")

# create list of companies to run regression for
Company_list <- unique(complete_Data$ID)

# run regression in loop and store estimates in complete_Data
for (i in Company_list) {
  s <- subset(complete_Data, ID == i)
  for (t in 1:(length(s$beta_monthly)-w+1)) {
    if(t <= 18){v <- 18} else {v <- 36}
    if(t <= 18){ss <- s[1:(v+t-1),]} else {ss <- s[(1+t-19):(v+t-19),]}
    v <- 36
    ss <- filter(ss, !is.na(ExcessReturn))
    if(nrow(ss)< w){next()}
    tryCatch({
      reggg <- lm(formula = ExcessReturn ~ EPU_Index, data = ss)
      s <- s[order(as.Date(s$Date, format="%d/%m/%Y")),]
      s$beta_monthly[v-19+t] <- coefficients(reggg)[2]
    }, error =function(e) {
      s <- s[order(as.Date(s$Date, format="%d/%m/%Y")),]
      s$beta_monthly[v-19+t] <- NA
    })}
  complete_Data <- merge(complete_Data, subset(s,select =
                          c("RowNr", "beta_monthly")), by = "RowNr", all = TRUE)
  complete_Data$beta_monthly <- coalesce(complete_Data$beta_monthly.x,
                                         complete_Data$beta_monthly.y)
  complete_Data <- subset(complete_Data, select = -c(beta_monthly.x,
                                                     beta_monthly.y))}

#remove observations without excess return
Comp_beta <- filter(complete_Data, !is.na(ExcessReturn))
Comp_beta <- filter(Comp_beta, !is.na(beta_monthly))

# rename dataframe and save
Comp_beta_Print <- Comp_beta
save(Comp_beta_Print, file = "Comp_beta_Print.Rdata")


####----------------------------------------------------------------####
####          Section 7 - Textual analysis of annual reports         ####
####----------------------------------------------------------------####

# define filepath to annual reports folder
filepath <- "C:/Users/jenst/OneDrive/Documents/Master Thesis/Reports1"

# read in annual reports
allfiles1 <- readtext(filepath, text_field = "texts")

# remove whitespace
allfiles <- stri_trim(allfiles1)

# transform to lower case letters
allfiles <- stri_trans_tolower(allfiles)

#remove numbers and punctuation
allfiles <- stringi::stri_replace_all_regex(allfiles, "\\d", "")
```

```
allfiles <- stringi::stri_replace_all_regex(allfiles, "[\\p{p}\\p{s}]", "")

# concatenate word tokens which belong together
for (j in seq(allfiles))
{
  allfiles[[j]] <- gsub("central bank*", "central_bank", allfiles[[j]])
  allfiles[[j]] <- gsub("norges bank", "norges_bank", allfiles[[j]])
}

# split in tokens(words)
filetokens <- tokens(allfiles)

# load norwegian and english stopwords like "I", "me", "det", "og", etc.
NORstopword <- stopwords("norwegian")
ENGstopword <- stopwords("english")

# remove stopwords
filetokens <- tokens_remove(filetokens, NORstopword)
filetokens <- tokens_remove(filetokens, ENGstopword)

# create dictionary of words to count
dict <- dictionary(list(norges_bank = "norges_bank",
              central_bank = "central_bank*", government = "government*",
              ministry = c("ministry","ministries"),
              regulation = "regulation*", minister = "minister*",
              directive = "directive*",parliament = "parliament",
              sentralbank = "sentralbank*", regjering = "regjering*",
              departement = "departement*", regulering = "regulering*",
              minister = "minster*",direktiv = "direktiv*",
              storting = "storting*", myndigheter = "myndighete*"))

# create document-term-matrix (dtm)
dtm <- dfm(filetokens)

# count occasions of each dictionary word in each file in the dtm
dict_dtm <- dfm_lookup(dtm, dict, nomatch = "_unmatched")

# convert dtm to dataframe and merge in doc_id
dict_dtm_frame <- as.data.frame(dict_dtm)
dict_dtm_frame$doc_id <- allfiles1$doc_id

# sum count of words matching those from the dictionary
dict_dtm_frame$match <- rowSums(dict_dtm_frame[,2:16])

# sum matching and unmatching words
dict_dtm_frame$total <- dict_dtm_frame$`_unmatched` +
  dict_dtm_frame$match

# calculate relative number of matching words
dict_dtm_frame$frequency <- dict_dtm_frame$match/dict_dtm_frame$total

# load data
load("TestAssetData.Rdata")

# rename
dtm <- dict_dtm_frame

# split doc_id strings to get company name and year in separate columns
dtm <- data.frame(dtm,do.call(rbind,str_split(dtm$doc_id,"_")))
dtm$Year <- str_split_fixed(dtm$X2, ".pdf", 4)
dtm$company <- dtm$X1
```

```
dtm <- subset(dtm, select = -c(X1, X2))

# remove Gyldendal ASA 2012 and 2013 because they are scanned and
# words cannot be counted
dtm <- dtm[-c(781, 782), ]

# standardize values
sd <- aggregate(dtm$frequency, by=list(dtm$company), FUN=sd)
colnames(sd)[1:2] <- c("company", "sd")
dtm <- merge(dtm, sd, by = "company")
dtm$adjFreq <- dtm$frequency/dtm$sd

# keep only first level of Year column
dtm$Year <- (dtm$Year[,1])
#keep first four characters of Year columns
dtm$Year <- substr(dtm$Year, 0,4)
# change to numeric
dtm$Year <- as.numeric(as.character(dtm$Year))
# add plus 1 to year
dtm$Year <- (dtm$Year+1)

# expand each observation to compare yearly portfolio updates with
# monthly company returns
Month = expand.grid(Year = unique(dtm$Year), Month = 1:12)
dtm <- left_join(dtm, Month, by = "Year")

# create a column to denote what monthly returns to sum from june to june
dtm$return_Year = ifelse(dtm$Month > 6, dtm$Year, dtm$Year+1)

# adjust names
dtm$company <- stringi::stri_replace_all_regex(dtm$company,
               "AMERICAN SHIPPING COMPANY", "AMERICAN SHIPPING CO ASA")
dtm$company <- stringi::stri_replace_all_regex(dtm$company,
               "KONGSBERG GRUPPEN", "KONGSBERG GRUPPEN ASA")
dtm$company <- stringi::stri_replace_all_regex(dtm$company,
               "LER??Y SEAFOOD GROUP ASA", "LEROY SEAFOOD GROUP ASA")
dtm$company <- stringi::stri_replace_all_regex(dtm$company,
               "NORDIC SEMICONDUCTOR ASA", "NORDIC SEMICONDUCTOR")
dtm$company <- stringi::stri_replace_all_regex(dtm$company,
               "ODDFJELL SE", "ODFJELL SE")

# subset dataframe and omit NAs
Yearly_returns <- subset(TestAssetData, select = c("return_Year",
                  "company", "ExcessReturn","Month", "ID", "Year"))
Yearly_returns <- na.omit(Yearly_returns)

# calculate average return in each year to be able to estimate a yearly
# beta
Yearly_returns <- aggregate(Yearly_returns$ExcessReturn,
                by=list(Yearly_returns$ID, Yearly_returns$Year), FUN=mean)
colnames(Yearly_returns)[1:3] <- c("ID", "Year", "ExcessReturn")

#
names <- subset(TestAssetData, select = c("company", "ID"))
names <- names %>% distinct(ID, .keep_all = TRUE)
Yearly_returns <- merge(Yearly_returns, names, by = "ID")

#remove "/" in company names
Yearly_returns$company <-
  stringi::stri_replace_all_regex(Yearly_returns$company, "/", "")
```

```r
# merge and subset dataframes
Yearly_returns <- merge(dtm, Yearly_returns, by =
                          c("company", "Year"), all = TRUE)
Yearly_returns <- subset(Yearly_returns, Year > 2008)
Yearly_returns <- na.omit(Yearly_returns)

# remove companies with unnatural observations
Yearly_returns <- subset(Yearly_returns, company != "AKASTOR ASA")
Yearly_returns <- subset(Yearly_returns, company != "HUNTER GROUP ASA")
Yearly_returns <- subset(Yearly_returns, company != "COPEINCA ASA")

# subset dataframe and keep unique rows based on ID and Year
Returns <- subset(Yearly_returns, select =
                    c("company", "Year", "doc_id", "adjFreq", "Month",
                      "return_Year","ID", "ExcessReturn"))
Returns <- Returns %>% relocate(ID, Year)
Returns <- Returns[!duplicated(Returns[1:2]),]

# create column to store EPU betas
Returns$beta_yearly <- NA

# No. of observations in rolling regression
v <- 12

# No of observations of each stock requred in rolling regression
w <- 6

# set up dataframe
complete_Data <- Returns
complete_Data <- complete(complete_Data, ID, Year)
complete_Data <- tibble::rowid_to_column(complete_Data, "RowNr")

# create list of companies to run regression for
Company_list <- unique(complete_Data$ID)

# run regression in loop and store estimates in complete_Data
for (i in Company_list) {
  s <- subset(complete_Data, ID == i)
  for (t in 1:(length(s$beta_yearly)-w+1)) {
    if(t <= 6){v <- 6} else {v <- 11}
    if(t <= 6){ss <- s[1:(v+t-1),]} else {ss <- s[(1+t-7):(v+t-7),]}
    v <- 12
    ss <- filter(ss, !is.na(ExcessReturn))
    if(nrow(ss)< w){next()}
    tryCatch({
      reggg <- lm(formula = ExcessReturn ~ adjFreq, data = ss)
      s <- s[order(s$Year),]
      s$beta_yearly[v-7+t] <- coefficients(reggg)[2]
    }, error =function(e) {
      s <- s[order(s$Year),]
      s$beta_yearly[v-7+t] <- NA
    })
  }
  complete_Data <- merge(complete_Data, subset(s,select =
                      c("RowNr", "beta_yearly")), by = "RowNr", all = TRUE)
  complete_Data$beta_yearly <- coalesce(complete_Data$beta_yearly.x,
                                        complete_Data$beta_yearly.y)
  complete_Data <- subset(complete_Data, select = -c(beta_yearly.x,
                                                      beta_yearly.y))
}
```

```
#remove observations without excess return, subset and rename columns
Comp_beta <- filter(complete_Data, !is.na(ExcessReturn))
Comp_beta <- filter(Comp_beta, !is.na(beta_yearly))
Comp_beta <- subset(Comp_beta, select = c("ID", "Year","beta_yearly",
                                          "doc_id"))
colnames(Comp_beta)[3] <- ("beta_monthly")

# expand dataframe from yearly to monthly, and merge in other sorting
# variables
Month = expand.grid(Year = unique(Comp_beta$Year), Month = 1:12)
Comp_beta <- left_join(Comp_beta, Month, by = "Year")
Comp_beta <- merge(Comp_beta, TestAssetData, by =
                     c("ID", "Year", "Month"))

# rename and save dataframe
Comp_beta_TextAnalysis <- Comp_beta
save(Comp_beta_TextAnalysis, file = "Comp_beta_TextAnalysis.Rdata")



####------------------------------------------------------------------####
####                    Section 8 - Create factors                    ####
####------------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Online", "Google" or
### "TextAnalysis".

# load dataset
load("Comp_beta_Print.Rdata")

# rename and subset dataframe
factor_data <- Comp_beta_Print
factor_data1 <- subset(factor_data, Month == 6)

# cut in 0.3 and 0.7 quantiles to create groups for construction
# of factors
factor_data1 <- ddply(factor_data1, .(Year), mutate,
                     SIZE_gr = ntile(SIZE, 2),
                     BM_gr = cut(BM, quantile(BM, c(0, 0.3, 0.7, 1)),
                                 labels = c(1,2,3), include.lowest = T),
                     OP_gr = cut(OP, quantile(OP, c(0, 0.3, 0.7, 1)),
                                 labels = c(1,2,3), include.lowest = T),
                     INVEST_gr = cut(INVEST, quantile(INVEST,
                                  c(0, 0.3, 0.7, 1)),
                                  labels = c(1,2,3), include.lowest = T),
                     EPU_gr = cut(beta_monthly, quantile(beta_monthly,
                                  c(0, 0.3, 0.7, 1)),
                                  labels = c(1,2,3), include.lowest = T))

# assign stocks to factor mimicking groups
factor_data1 <- transform(factor_data1 , SIZE_BM = paste0(SIZE_gr, BM_gr),
                         SIZE_INVEST = paste0(SIZE_gr, INVEST_gr),
                         SIZE_OP = paste0(SIZE_gr, OP_gr),
                         SIZE_EPU = paste0(SIZE_gr, EPU_gr))

# subset dataframe
factor_data1 <- subset(factor_data1, select =
                         -c(return_Year, Month, ExcessReturn, MCAP))

# expand each observation to compare yearly portfolio updates with
# monthly company returns
```

```r
Month = expand.grid(Year = unique(factor_data1$Year), Month = 1:12)
factor_data1 <- left_join(factor_data1, Month, by = "Year")

# create a column to denote what monthly returns to sum from july to june
factor_data1$return_Year = ifelse(factor_data1$Month > 6,
                                  factor_data1$Year,factor_data1$Year+1)

# subset dataframe
factor_data <- subset(factor_data, select = c("ID", "return_Year",
                                  "Month", "MCAP", "ExcessReturn"))
factor_data <- merge(factor_data1,factor_data, by = c("return_Year",
                                                      "Month", "ID"))

# estimating return of factor mimicking portfolios at each time t
factor_data <- data.table(factor_data)
SMB_BM <- factor_data[, list(Return_SMB_BM =
                            weighted.mean(ExcessReturn, MCAP, na.rm = T),
                            MCAP = sum(MCAP)), by =
                        c("return_Year", "Month", "SIZE_BM", "SIZE_gr")]
SMB_OP <- factor_data[, list(Return_SMB_OP =
                            weighted.mean(ExcessReturn, MCAP, na.rm = T),
                            MCAP = sum(MCAP)),by =
                        c("return_Year", "Month", "SIZE_OP", "SIZE_gr")]
SMB_INVEST <- factor_data[, list(Return_SMB_INVEST =
                            weighted.mean(ExcessReturn, MCAP, na.rm = T),
                                MCAP = sum(MCAP)), by =
                        c("return_Year", "Month", "SIZE_INVEST", "SIZE_gr")]
SMB_EPU <- factor_data[, list(Return_SMB_EPU =
                            weighted.mean(ExcessReturn, MCAP, na.rm = T),
                             MCAP = sum(MCAP)), by =
                        c("return_Year", "Month", "SIZE_EPU", "SIZE_gr")]

# estimating each SMB-sort return
SMB_BM1 <- SMB_BM[, list(Return_SMB_BM =
                           weighted.mean(Return_SMB_BM, MCAP)),
                           by = c("return_Year", "Month", "SIZE_gr")]
SMB_OP1 <- SMB_OP[, list(Return_SMB_OP =
                           weighted.mean(Return_SMB_OP, MCAP)),
                           by = c("return_Year", "Month", "SIZE_gr")]
SMB_INVEST1 <- SMB_INVEST[, list(Return_SMB_INVEST =
                           weighted.mean(Return_SMB_INVEST, MCAP)),
                           by = c("return_Year", "Month", "SIZE_gr")]
SMB_EPU1 <- SMB_EPU[, list(Return_SMB_EPU =
                           weighted.mean(Return_SMB_EPU, MCAP)),
                           by = c("return_Year", "Month", "SIZE_gr")]

# order dataframes by return_Year, Month and descending SIZE_gr
SMB_BM1 <- SMB_BM1[order(return_Year, Month, -SIZE_gr),]
SMB_OP1 <- SMB_OP1[order(return_Year, Month, -SIZE_gr),]
SMB_INVEST1 <- SMB_INVEST1[order( return_Year, Month, -SIZE_gr),]
SMB_EPU1 <- SMB_EPU1[order(return_Year, Month, -SIZE_gr),]

# calculate difference in return between the two SIZE groups in each
# time period
SMB_BM2 <- SMB_BM1[, list(SMB_BM = diff(Return_SMB_BM)),
                   by = c("return_Year", "Month")]
SMB_OP2 <- SMB_OP1[, list(SMB_OP = diff(Return_SMB_OP)),
                   by = c("return_Year", "Month")]
SMB_INVEST2 <- SMB_INVEST1[, list(SMB_INVEST = diff(Return_SMB_INVEST)),
                           by = c("return_Year", "Month")]
SMB_EPU2 <- SMB_EPU1[, list(SMB_EPU = diff(Return_SMB_EPU)),
```

```
                     by = c("return_Year", "Month")]

# convert to dataframe
SMB_BM2 <- as.data.frame(SMB_BM2)
SMB_OP2 <- as.data.frame(SMB_OP2)
SMB_INVEST2 <- as.data.frame(SMB_INVEST2)
SMB_EPU2 <- as.data.frame(SMB_EPU2)

# merge dataframes
Factors <- merge(SMB_BM2, SMB_OP2, by = c("return_Year", "Month"))
Factors <- merge(Factors, SMB_INVEST2, by = c("return_Year", "Month"))
Factors <- merge(Factors, SMB_EPU2, by = c("return_Year", "Month"))

# calculate factor
Factors$SMB <- (Factors$SMB_BM + Factors$SMB_OP + Factors$SMB_INVEST +
                  Factors$SMB_EPU)/4

# subset dataframe
Factors <- subset(Factors, select = c("return_Year", "Month", "SMB"))

# Calculate HML, RMW, CMA and EPU
# estimate return of factor mimicking portfolios at each time t
SMB_BM4 <-  factor_data[, list(Return_SMB_BM =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T),
                              MCAP = sum(MCAP)), by = c("return_Year",
                                        "Month", "SIZE_BM", "BM_gr")]
SMB_OP4 <-  factor_data[, list(Return_SMB_OP =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T),
                              MCAP = sum(MCAP)),by = c("return_Year",
                                        "Month", "SIZE_OP", "OP_gr")]
SMB_INVEST4 <-  factor_data[, list(Return_SMB_INVEST =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T),
                              MCAP = sum(MCAP)), by = c("return_Year",
                                        "Month", "SIZE_INVEST", "INVEST_gr")]
SMB_EPU4 <-  factor_data[, list(Return_SMB_EPU =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T),
                              MCAP = sum(MCAP)), by = c("return_Year",
                                        "Month", "SIZE_EPU", "EPU_gr")]

# estimating returns
SMB_BM4 <- SMB_BM4[, list(Return_SMB_BM =
                                weighted.mean(Return_SMB_BM, MCAP)),
                                by = c("return_Year", "Month", "BM_gr")]
SMB_OP4 <- SMB_OP4[, list(Return_SMB_OP =
                                weighted.mean(Return_SMB_OP, MCAP)),
                               by = c("return_Year", "Month", "OP_gr")]
SMB_INVEST4 <- SMB_INVEST4[, list(Return_SMB_INVEST =
                                weighted.mean(Return_SMB_INVEST, MCAP)),
                                by = c("return_Year", "Month", "INVEST_gr")]
SMB_EPU4 <- SMB_EPU4[, list(Return_SMB_EPU =
                                weighted.mean(Return_SMB_EPU, MCAP)),
                                by = c("return_Year", "Month", "EPU_gr")]

# keep only low and high quantiles for each sort
SMB_BM4 <- subset(SMB_BM4, BM_gr != "2")
SMB_OP4 <- subset(SMB_OP4, OP_gr != "2")
SMB_INVEST4 <- subset(SMB_INVEST4, INVEST_gr != "2")
SMB_EPU4 <- subset(SMB_EPU4, EPU_gr != "2")

# order dataframes by return_Year, Month and descending sorting variable
SMB_BM4 <- SMB_BM4[order(return_Year, Month, -BM_gr),]
```

```
SMB_OP4 <- SMB_OP4[order(return_Year, Month, -OP_gr),]
SMB_INVEST4 <- SMB_INVEST4[order(return_Year, Month, -INVEST_gr),]
SMB_EPU4 <- SMB_EPU4[order(return_Year, Month, -EPU_gr),]

# calculate difference in return between the two sorting groups in
# each time period
SMB_BM4 <- SMB_BM4[, list(SMB_BM = diff(-Return_SMB_BM)),
                   by = c("return_Year", "Month")]
SMB_OP4 <- SMB_OP4[, list(SMB_OP = diff(-Return_SMB_OP)),
                   by = c("return_Year", "Month")]
SMB_INVEST4 <- SMB_INVEST4[, list(SMB_INVEST = diff(Return_SMB_INVEST)),
                           by = c("return_Year", "Month")]
SMB_EPU4 <- SMB_EPU4[, list(SMB_EPU = diff(Return_SMB_EPU)),
                     by = c("return_Year", "Month")]

# combine factors in one dataframe
Factors$HML <- SMB_BM4$SMB_BM
Factors$RMW <- SMB_OP4$SMB_OP
Factors$CMA <- SMB_INVEST4$SMB_INVEST
Factors$EPU <- SMB_EPU4$SMB_EPU

# construct market factor from test asset sample
load("stockdata_monthly.R")

# remove NAs
stockdata_monthly <-
  stockdata_monthly[!is.na(stockdata_monthly$ExcessReturn),]
stockdata_monthly <- stockdata_monthly[!is.na(stockdata_monthly$MCAP),]

# subset dataframe
Index_finalstock <- subset(stockdata_monthly,
                    select = c("Year", "Month", "MCAP", "ExcessReturn"))

# change format to data table and calculate market return
Index_finalstock <- data.table(Index_finalstock)
Index_finalstock <- Index_finalstock[, list(finalstock =
                                     weighted.mean(ExcessReturn, MCAP)),
                                     by = c("Year", "Month")]
# rename variable
Index_finalstock <- Index_finalstock %>%
  dplyr::rename(
     return_Year = Year)

Index_finalstock <- data.frame(Index_finalstock)

Factors <- merge(Factors, Index_finalstock,by = c("return_Year", "Month"))

save(Factors, file = "Factors.Rdata")
save(Factors, file = "Print_Factors.Rdata")
```

```r
####-----------------------------------------------------------------####
####              Section 9 - Calculate test asset returns            ####
####-----------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Online", "Google" or
### "TextAnalysis".

load("Print_Factors.Rdata")
load(file = "TestAssetData.Rdata")

# remove observations after 6/2019
Factors <- Factors[!(Factors$return_Year == 2019 & Factors$Month > 6), ]

# estimate return of each test asset
TestAssetData <- data.table(TestAssetData)
SIZE_BM_return <- TestAssetData[, list(Return_SIZE_BM =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T)),
                              by = c("return_Year", "Month", "SIZE_BM")]
SIZE_OP_return <- TestAssetData[, list(Return_SIZE_OP =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T)),
                              by = c("return_Year", "Month", "SIZE_OP")]
SIZE_INVEST_return <- TestAssetData[, list(Return_SIZE_INVEST =
                              weighted.mean(ExcessReturn, MCAP, na.rm = T)),
                              by = c("return_Year", "Month", "SIZE_INVEST")]

# merge factors with test asset data
SIZE_BM_return <- merge(SIZE_BM_return, Factors, by =
                          c("return_Year", "Month"))
SIZE_INVEST_return <- merge(SIZE_INVEST_return, Factors,
                            by = c("return_Year", "Month"))
SIZE_OP_return <- merge(SIZE_OP_return, Factors,
                        by = c("return_Year", "Month"))

# make complete dataframes with observations each month each year for all
# stocks from first to last observed time
Factors <- Factors %>% arrange(return_Year, Month)
mindate <- paste(min(Factors$return_Year),"-", Factors$Month[1],
                 "-","01", sep = "")
maxdate <- paste(max(Factors$return_Year),"-",
                 Factors$Month[nrow(Factors)], "-","01", sep = "")
dateseq <- as.data.frame(seq(as.Date(mindate),as.Date(maxdate),
                             by = "month"))
colnames(dateseq)[1] <- "Date"
dateseq <- dateseq %>%
  dplyr::mutate(return_Year = lubridate::year(Date),
                Month = lubridate::month(Date))
dateseq <- subset(dateseq, select = c("return_Year", "Month"))

# make complete dataframe
Test_Asset_gr <- unique(SIZE_BM_return[, c("SIZE_BM")])
Test_Asset_date <- merge(dateseq, Test_Asset_gr)
SIZE_BM_return <- merge(SIZE_BM_return, Test_Asset_date,
                        by = c("return_Year", "Month", "SIZE_BM"),
                        all = TRUE)

# make complete dataframe
Test_Asset_gr <- unique(SIZE_OP_return[, c("SIZE_OP")])
Test_Asset_date <- merge(dateseq, Test_Asset_gr)
SIZE_OP_return <- merge(SIZE_OP_return, Test_Asset_date,
                        by = c("return_Year", "Month", "SIZE_OP"),
                        all = TRUE)
```

```
# make complete dataframe
Test_Asset_gr <- unique(SIZE_INVEST_return[, c("SIZE_INVEST")])
Test_Asset_date <- merge(dateseq, Test_Asset_gr)
SIZE_INVEST_return <- merge(SIZE_INVEST_return, Test_Asset_date,
                            by = c("return_Year", "Month", "SIZE_INVEST"),
                            all = TRUE)
SIZE_INVEST_return[is.na(SIZE_INVEST_return)] <- 0

# save test asset return dataframes
save(SIZE_BM_return, file = "Print_SIZE_BM_return.Rdata")
save(SIZE_OP_return, file = "Print_SIZE_OP_return.Rdata")
save(SIZE_INVEST_return, file = "Print_SIZE_INVEST_return.Rdata")



####----------------------------------------------------------------####
####             Section 10 - Fama Macbeth regressions              ####
####----------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Online", "Google" or
### "TextAnalysis" and to replace "SIZE_BM" with "SIZE_OP" and
### "SIZE_INVEST". Then run code for each measure and each double sort.

# load file
load(file = "Print_SIZE_BM_return.Rdata")

# create list of test assets to run regression for
Test_Asset_gr <- unique(SIZE_BM_return$SIZE_BM)

# create data frames to store regression results
SIZE_BM_Betas_CAPM <- as.data.frame(matrix(NA, nrow = 9, ncol = 3))
SIZE_BM_ALL_CAPM <- as.data.frame(matrix(NA, nrow = 27, ncol = 5))

# set counting variable to zero
w <-0

# for loop to run Fama Macbeth step 1 regression for each test asset,
# and store result in dataframe
for (i in Test_Asset_gr) {
  s <- subset(SIZE_BM_return, SIZE_BM == i)
  w <- w+1
  reg <- lm(formula = Return_SIZE_BM ~ finalstock + EPU, data = s)
  SIZE_BM_Betas_CAPM[w,1:3] <- reg$coefficients
  SIZE_BM_ALL_CAPM[w,1:3] <- reg$coefficients
  SIZE_BM_ALL_CAPM[w,4] <- i
  SIZE_BM_ALL_CAPM[w,5] <- "coefficient"
  SIZE_BM_ALL_CAPM[(w+9),1:3] <- summary(reg)$coefficients[,3]
  SIZE_BM_ALL_CAPM[(w+9),4] <- i
  SIZE_BM_ALL_CAPM[(w+9),5] <- "t-stat"
  SIZE_BM_ALL_CAPM[(w+18),1:3] <- summary(reg)$coefficients[,4]
  SIZE_BM_ALL_CAPM[(w+18),4] <- i
  SIZE_BM_ALL_CAPM[(w+18),5] <- "P-value"
}

# rename columns
colnames(SIZE_BM_Betas_CAPM) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_CAPM) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_CAPM)[4:5] <-c("TestAsset", "Value")

# define column
```

```
SIZE_BM_Betas_CAPM$SIZE_BM <- Test_Asset_gr

# subset, merge and omit NAs
SIZE_BM_m <- subset(SIZE_BM_return, select =
                    c("return_Year", "Month", "SIZE_BM", "Return_SIZE_BM"))
SIZE_BM_Betas <- merge(SIZE_BM_Betas_CAPM, SIZE_BM_m, by = "SIZE_BM")
SIZE_BM_Betas <- na.omit(SIZE_BM_Betas)

# subset and save dataframes
SIZE_BM_tablestep1_CAPM <- subset(SIZE_BM_ALL_CAPM, select =
                                  c("EPU", "TestAsset"))
save(SIZE_BM_tablestep1_CAPM,file = "Print_SIZE_BM_tablestep1_CAPM.Rdata")
save(SIZE_BM_ALL_CAPM, file = "Print_SIZE_BM_ALL_CAPM.Rdata")


# create unique date indicator
SIZE_BM_Betas$Dates <- paste(SIZE_BM_Betas$return_Year,
                             SIZE_BM_Betas$Month, sep = "")

# order dataframe
SIZE_BM_Betas <- SIZE_BM_Betas %>% arrange(return_Year, Month)

# make dataframe with unique date indicators for each time period
dates <- unique(SIZE_BM_Betas$Dates)

# create and edit dataframe to store estimated premiums in
SIZE_BM_Premiums_CAPM <- subset(SIZE_BM_Betas, select =
                                -c(Return_SIZE_BM, SIZE_BM))
SIZE_BM_Premiums_CAPM =
  SIZE_BM_Premiums_CAPM[!duplicated(SIZE_BM_Premiums_CAPM$Dates),]
SIZE_BM_Premiums_CAPM <- SIZE_BM_Premiums_CAPM %>%
  arrange(return_Year, Month)
SIZE_BM_Premiums_CAPM[,1:3] <- NA

# set counting variable to zero
w <- 0

# for loop to run Fama Macbeth step 2 regression for each time period,
# and store result in dataframe
for (i in dates) {
  w <- w+1
  s <- subset(SIZE_BM_Betas, Dates == i)
  reg <- lm(formula = Return_SIZE_BM ~ finalstock + EPU, data = s)
  SIZE_BM_Premiums_CAPM[w, 1:3]<- reg$coefficients
}

# run t-test for each factor
x <- lapply(SIZE_BM_Premiums_CAPM[,2:3], t.test)

# create dataframe with mean, p-value and t-statistic for each factor
statpremium <-
  t(data.frame(mean = sapply(x, getElement, name = "estimate"),
               p.value = sapply(x, getElement, name = "p.value"),
               tstat = sapply(x, getElement, name = "statistic")))
statpremium <- data.frame(statpremium)


# set column names
statpremium <- setNames(statpremium, names(SIZE_BM_Premiums_CAPM[,2:3]))

SIZE_BM_ALL_CAPM_Step2 <- statpremium
```

```
SIZE_BM_tablestep2_CAPM <- subset(statpremium, select = "EPU")

save(SIZE_BM_tablestep2_CAPM, file = "Print_SIZE_BM_tablestep2_CAPM.Rdata")
save(SIZE_BM_ALL_CAPM_Step2, file = "Print_SIZE_BM_ALL_CAPM_Step2.Rdata")



####------------------------------------------------------------------####
####           Section 11 - Fama Macbeth regressions Three-factor       ####
####------------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Online", "Google" or
### "TextAnalysis" and to replace "SIZE_BM" with "SIZE_OP" and
### "SIZE_INVEST". Then run code for each measure and each double sort.

# load file
load(file = "Print_SIZE_BM_return.Rdata")

# create list of test assets to run regression for
Test_Asset_gr <- unique(SIZE_BM_return$SIZE_BM)

# create data frames to store regression results
SIZE_BM_Betas_3F <- as.data.frame(matrix(NA, nrow = 9, ncol = 5))
SIZE_BM_ALL_3F <- as.data.frame(matrix(NA, nrow = 27, ncol = 7))

# set counting variable to zero
w <-0

# for loop to run Fama Macbeth step 1 regression for each test asset,
# and store result in dataframe
for (i in Test_Asset_gr) {
  s <- subset(SIZE_BM_return, SIZE_BM == i)
  w <- w+1
  reg <- lm(formula = Return_SIZE_BM ~ finalstock + SMB + HML + EPU,
            data = s)
  SIZE_BM_Betas_3F[w,1:5] <- reg$coefficients
  SIZE_BM_ALL_3F[w,1:5] <- reg$coefficients
  SIZE_BM_ALL_3F[w,6] <- i
  SIZE_BM_ALL_3F[w,7] <- "coefficient"
  SIZE_BM_ALL_3F[(w+9),1:5] <- summary(reg)$coefficients[,3]
  SIZE_BM_ALL_3F[(w+9),6] <- i
  SIZE_BM_ALL_3F[(w+9),7] <- "t-stat"
  SIZE_BM_ALL_3F[(w+18),1:5] <- summary(reg)$coefficients[,4]
  SIZE_BM_ALL_3F[(w+18),6] <- i
  SIZE_BM_ALL_3F[(w+18),7] <- "P-value"
}

# rename columns
colnames(SIZE_BM_Betas_3F) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_3F) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_3F)[6:7] <-c("TestAsset", "Value")

# define column
SIZE_BM_Betas_3F$SIZE_BM <- Test_Asset_gr

# subset, merge and omit NAs
SIZE_BM_m <- subset(SIZE_BM_return, select =
                    c("return_Year", "Month", "SIZE_BM", "Return_SIZE_BM"))
SIZE_BM_Betas <- merge(SIZE_BM_Betas_3F, SIZE_BM_m, by = "SIZE_BM")
SIZE_BM_Betas <- na.omit(SIZE_BM_Betas)
```

```r
# subset and save dataframes
SIZE_BM_tablestep1_3F <- subset(SIZE_BM_ALL_3F, select =
                                  c("EPU", "TestAsset"))
save(SIZE_BM_tablestep1_3F, file = "Print_SIZE_BM_tablestep1_3F.Rdata")
save(SIZE_BM_ALL_3F, file = "Print_SIZE_BM_ALL_3F.Rdata")


# create unique date indicator
SIZE_BM_Betas$Dates <-
  paste(SIZE_BM_Betas$return_Year, SIZE_BM_Betas$Month, sep = "")

# order dataframe
SIZE_BM_Betas <- SIZE_BM_Betas %>% arrange(return_Year, Month)

# make dataframe with unique date indicators for each time period
dates <- unique(SIZE_BM_Betas$Dates)

# create and edit dataframe to store estimated premiums in
SIZE_BM_Premiums_3F <- subset(SIZE_BM_Betas, select =
                                -c(Return_SIZE_BM, SIZE_BM))
SIZE_BM_Premiums_3F =
  SIZE_BM_Premiums_3F[!duplicated(SIZE_BM_Premiums_3F$Dates),]
SIZE_BM_Premiums_3F <- SIZE_BM_Premiums_3F %>% arrange(return_Year, Month)
SIZE_BM_Premiums_3F[,1:5] <- NA

# set counting variable to zero
w <- 0

# for loop to run Fama Macbeth step 2 regression for each time period,
# and store result in dataframe
for (i in dates) {
  w <- w+1
  s <- subset(SIZE_BM_Betas, Dates == i)
  reg <- lm(formula = Return_SIZE_BM ~ finalstock + SMB + HML + EPU,
            data = s)
  SIZE_BM_Premiums_3F[w, 1:5]<- reg$coefficients
}

# run t-test for each factor
x <- lapply(SIZE_BM_Premiums_3F[,2:5], t.test)

# create dataframe with mean, p-value and t-statistic for each factor
statpremium <-
  t(data.frame(mean = sapply(x, getElement, name = "estimate"),
               p.value = sapply(x, getElement, name = "p.value"),
               tstat = sapply(x, getElement, name = "statistic")))
statpremium <- data.frame(statpremium)

# set column names
statpremium <- setNames(statpremium, names(SIZE_BM_Premiums_3F[,2:5]))

SIZE_BM_ALL_3F_Step2 <- statpremium
SIZE_BM_tablestep2_3F <- subset(statpremium, select = "EPU")

save(SIZE_BM_tablestep2_3F, file = "Print_SIZE_BM_tablestep2_3F.Rdata")
save(SIZE_BM_ALL_3F_Step2, file = "Print_SIZE_BM_ALL_3F_Step2.Rdata")
```

```
####---------------------------------------------------------------------####
####          Section 12 - Fama Macbeth regressions Five-factor          ####
####---------------------------------------------------------------------####

### Use "Replace all" to replace "Print" with "Online", "Google" or
### "TextAnalysis" and to replace "SIZE_BM" with "SIZE_OP" and
### "SIZE_INVEST". Then run code for each measure and each double sort.

# load file
load(file = "Print_SIZE_BM_return.Rdata")

# create list of test assets to run regression for
Test_Asset_gr <- unique(SIZE_BM_return$SIZE_BM)

# create data frames to store regression results
SIZE_BM_Betas_5F <- as.data.frame(matrix(NA, nrow = 9, ncol = 7))
SIZE_BM_ALL_5F <- as.data.frame(matrix(NA, nrow = 27, ncol = 9))

# set counting variable to zero
w <-0

# for loop to run Fama Macbeth step 1 regression for each test asset,
# and store result in dataframe
for (i in Test_Asset_gr) {
  s <- subset(SIZE_BM_return, SIZE_BM == i)
  w <- w+1
  reg <- lm(formula =
              Return_SIZE_BM ~ finalstock + SMB + HML + RMW + CMA + EPU,
            data = s)
  SIZE_BM_Betas_5F[w,1:7] <- reg$coefficients
  SIZE_BM_ALL_5F[w,1:7] <- reg$coefficients
  SIZE_BM_ALL_5F[w,8] <- i
  SIZE_BM_ALL_5F[w,9] <- "coefficient"
  SIZE_BM_ALL_5F[(w+9),1:7] <- summary(reg)$coefficients[,3]
  SIZE_BM_ALL_5F[(w+9),8] <- i
  SIZE_BM_ALL_5F[(w+9),9] <- "t-stat"
  SIZE_BM_ALL_5F[(w+18),1:7] <- summary(reg)$coefficients[,4]
  SIZE_BM_ALL_5F[(w+18),8] <- i
  SIZE_BM_ALL_5F[(w+18),9] <- "P-value"
}

# rename columns
colnames(SIZE_BM_Betas_5F) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_5F) <- names(reg$coefficients)
colnames(SIZE_BM_ALL_5F)[8:9] <-c("TestAsset", "Value")

# define column
SIZE_BM_Betas_5F$SIZE_BM <- Test_Asset_gr

# subset, merge and omit NAs
SIZE_BM_m <- subset(SIZE_BM_return, select =
                      c("return_Year", "Month", "SIZE_BM", "Return_SIZE_BM"))
SIZE_BM_Betas <- merge(SIZE_BM_Betas_5F, SIZE_BM_m, by = "SIZE_BM")
SIZE_BM_Betas <- na.omit(SIZE_BM_Betas)

# subset and save dataframes
SIZE_BM_tablestep1_5F <- subset(SIZE_BM_ALL_5F, select =
                                  c("EPU", "TestAsset"))
save(SIZE_BM_tablestep1_5F, file = "Print_SIZE_BM_tablestep1_5F.Rdata")
save(SIZE_BM_ALL_5F, file = "Print_SIZE_BM_ALL_5F.Rdata")
```

```r
# create unique date indicator
SIZE_BM_Betas$Dates <- paste(SIZE_BM_Betas$return_Year,
                             SIZE_BM_Betas$Month, sep = "")

# order dataframe
SIZE_BM_Betas <- SIZE_BM_Betas %>% arrange(return_Year, Month)

# make dataframe with unique date indicators for each time period
dates <- unique(SIZE_BM_Betas$Dates)

# create and edit dataframe to store estimated premiums in
SIZE_BM_Premiums_5F <- subset(SIZE_BM_Betas, select =
                                -c(Return_SIZE_BM, SIZE_BM))
SIZE_BM_Premiums_5F =
  SIZE_BM_Premiums_5F[!duplicated(SIZE_BM_Premiums_5F$Dates),]
SIZE_BM_Premiums_5F <- SIZE_BM_Premiums_5F %>% arrange(return_Year, Month)
SIZE_BM_Premiums_5F[,1:7] <- NA

# set counting variable to zero
w <- 0

# for loop to run Fama Macbeth step 2 regression for each time period,
# and store result in dataframe
for (i in dates) {
  w <- w+1
  s <- subset(SIZE_BM_Betas, Dates == i)
  reg <- lm(formula =
              Return_SIZE_BM ~ finalstock + SMB + HML + RMW + CMA + EPU,
            data = s)
  SIZE_BM_Premiums_5F[w, 1:7]<- reg$coefficients
}

# run t-test for each factor
x <- lapply(SIZE_BM_Premiums_5F[,2:7], t.test)

# create dataframe with mean, p-value and t-statistic for each factor
statpremium <-
  t(data.frame(mean = sapply(x, getElement, name = "estimate"),
               p.value = sapply(x, getElement, name = "p.value"),
               tstat = sapply(x, getElement, name = "statistic")))
statpremium <- data.frame(statpremium)

# set column names
statpremium <- setNames(statpremium, names(SIZE_BM_Premiums_5F[,2:7]))

SIZE_BM_ALL_5F_Step2 <- statpremium
SIZE_BM_tablestep2_5F <- subset(statpremium, select = "EPU")

save(SIZE_BM_tablestep2_5F, file = "Print_SIZE_BM_tablestep2_5F.Rdata")
save(SIZE_BM_ALL_5F_Step2, file = "Print_SIZE_BM_ALL_5F_Step2.Rdata")
```