Master Thesis

Financial Economics

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Natural Hazard Risk and the Effect on Future Stock Returns

Name Student: Joost Hendrik Albert Schipper
Student ID Number: 384171

Supervisor: dr. E Smajlbegovic
Second Assessor: dr. X Ma

30.04.2020

Table of Contents

Abstract

This research studies the effect of regional natural hazard risk to expected rates of return. I construct a novel proxy to measure regional natural hazard risk for states in the U.S. To construct firm-specific natural hazard risk, I identify all U.S. states that are economically relevant to firms and link this to the state-specific risk. I find some evidence that regional natural hazard risk can positively predict cross-sectional stock returns, but the explanatory power is limited.

1 Introduction

Climate change causes a significant increase in the amount of natural disasters (Coronese et al., 2019). Besides the frequency of natural disasters, the intensity and the economic damage of natural disasters are increasing (Stéphane, 2014). The severe effects of natural disasters are not to be underestimated with damages worldwide surpassing 100 billion US dollars in 2018 (Statista, 2019).

Even though several studies research the effect of natural disasters on the economy, literature does not agree on the impact on the economy. On the one hand researchers find short term GDP growth as a consequence of natural disasters (Albala-Bertrand, 1993; Caballeros Otero & Zapata Marti, 1995; Dacy & Kunreuther, 1969). On the other hand, a more recent strand of literature argues that natural disasters impact the GDP negatively (Hochrainer, 2009; Noy & Nualsri, 2011; Raddatz, 2009; Strobl, 2008; Xiao, 2011).

In related research, Tavor and Teitler-Regev (2019) find that natural disasters have a long lasting negative effect on stock prices. Wang and Kutan (2013) and Worthington (2008) find that natural disasters have no impact on the aggregate stock market. This is in line with Strobl (2011) who finds that the impact of natural disasters is significant at local level, but diversified away at national level. Despite a great deal of research to the effects of specific natural disasters, to my knowledge, the risk of natural disasters has yet to be used in explaining future stock returns.

Driven by the event-study based evidence on the effect of natural disasters to stock prices, in this paper, I research the potential link between natural hazard risk as a macroeconomic variable and the predictability of future stock returns. The main conjecture is that the demand for a firm’s stocks changes over time as the company’s specific natural hazard risk changes. If these demands shifts are systematic, they could impact current prices and future stock returns. This research builds on research of West and Lenze (1994) who argue that effects and consequences of natural disasters are mainly local. To incorporate the local risk to the U.S. stocks, I break down the U.S. market into states and obtain the state-specific natural hazard risk. By determining which states are economically relevant to firms, I can match the natural hazard risk to firms and obtain firm-specific macroeconomic risk figures.

To identify the state-specific natural hazard risk I construct a novel natural hazard risk proxy. Based on previous natural hazard literature of amongst others Birkmann (2007), the novel measure takes both regional exposure and vulnerability to natural disasters into account. The measure depends on historic events to calculate monthly expected exposure and current state-specific social economic variables for the vulnerability calculation.

To determine the economically relevant states for a company, I use a firm-specific measure constructed by Smajlbegovic (2019) in which each state is assigned a weight between 0 and 1. This weight is based on the share of citations of the specific state in an annual report in a certain year (i.e. the number of citations of a state divided by the total number of state citations). Weighting the monthly natural hazard risk for each state with the corresponding citation shares enables me to construct the firm-specific natural hazard risk proxy. The new proxy allows me to test whether natural hazard risk is able to predict the cross section of individual stock returns. I hypothesize that investors require compensation for higher natural hazard risk. More specific, if a company has higher natural hazard risk, I expect an increase in the cross section of individual excess stock returns.

I research the hypothesis with Fama and MacBeth (1973) regressions and quintile portfolio sorts based on natural hazard risk. From the FM regressions I find that natural hazard risk positively predicts differences in the cross section of excess stock returns when controlling for standard cross-sectional variables. The effect is robust on a 5% significance level to alternative explanations such as industry momentum (Moskowitz & Grinblatt, 1999), geographic momentum (Parsons, 2016) and the geographic dispersion of firms (García & Norli, 2012). Neither is the effect caused by comovement of excess returns of firms headquartered in the same state (Pirinsky & Wang, 2006). The results of the quintile portfolio sorts indicate that firms in the highest natural hazard risk portfolio outperform the lowest portfolio by, on average, 25 basis points per month. The corresponding t – statistics however are not significant by any standard confidence levels. When controlling for common risk factors, the portfolio sorts show that firms in the top portfolio of natural hazard risk are positively related to the market beta and the size factor, while the exposure to the value and momentum factors is strongly negative.

When adding the profitability and investment factors as controlling variables, the t – statistic of the excess returns becomes significant due to the negative exposure to the profitability factor. When controlling for common risk factors and standard control variables the explanatory power of natural hazard risk does not exceed the proposed t – statistic of 3.0. Therefore natural hazard risk cannot irrefutably be considered a factor in explaining cross-sectional results (Harvey, Liu, & Zhu, 2016).

Finally, to explore whether another explanation can be found for the results, I test whether the cross-sectional return predictability is associated with difficulty to arbitrage stocks. This might indicate that the predictability is due to mispricing based on (to me unknown) regional information.

The remainder of this research is structured as follows. Section 2 describes concepts and theories used for answering the research question and explains the position of this paper in relation to existing literature. Section 3 describes the data set and the used methodology. In Section 4 I provide the empirical results from the main analysis and in Section 5 I include corresponding robustness checks and additional insights. Finally, Section 6 summarizes and concludes the research.

2 Theoretical Background

This section describes and evaluates theories used to research the effect of regional natural hazard risk to stock returns. I conjecture that natural hazard risk is a factor to determine future stock returns. To examine this, first I discuss literature on measuring natural hazard risk. The following subsection describes the literature on the segmented market in the United States and economic relevance of a region for stock returns. Thereafter I describe the hypothesized relation between regional natural hazard risk and stock returns. Finally, I analyze the models chosen to answer the research question.

2.1 Measuring Natural Hazard Risk

In recent years, multiple initiatives have been taken to measure natural hazard risk at a local or global level. The goal of most of this research is to raise political awareness and provide an overview for efficient risk reduction. In 2003, the UN stressed that natural hazards become disasters when people's lives and livelihoods are affected (Annan, 2003). This contrasts to the literature on natural disasters of the past. Natural disaster research was primarily aimed at physical occurrences and did not encompass the susceptibility of people and communities or their abilities to deal with possible damages (Lewis, 1999). Following more recent literature, I consider natural disaster risk as a multiplication of exposure risk and vulnerability risk (Birkmann, 2006). Exposure is the relative number of people in a region who are exposed to natural hazards in a certain time period. Vulnerability relates to the physical, economic, social and environmental factors which make a region able to cope with the negative impact of natural disasters (Birkmann, 2007; Velásquez, Cardona, Carreño, & Barbat, 2014).

To capture the relevant exposure risk to natural hazards of a state, first I determine relevant natural disasters. I chose the disaster types based on the frequency of their occurrence since 1960 and the impact of such a disaster. In line with recent literature on natural disaster risk measurement I consider earthquakes, cyclone surges, cyclone winds and flooding (Birkmann, 2006; UNU-EHS, 2018). Natural hazards like drought and volcanic activity have been discarded because they both occur in less than 2% of the cases and there is little data available on the impact of these disasters in the United States. Wildfires have been discarded because of lack of data about the affected area and their relatively low impact on society. The WorldRiskIndex also takes sea-level rise into account (2018). I chose not to add sea-level rise to the exposure risk because a lot of the research about sea-level rise is very recent. Besides, it is a very gradual process, expected to take over a hundred years and it has not yet occurred, which makes it less suitable for this research. Even after determining the relevant disaster types, there are multiple ways to calculate the exposure to natural disasters. For example: The affected area of a state, the amount of casualties in a state or the number of people affected per state. Considering the large distinction of population density in the U.S. it would not make sense to choose the affected area. Following leading reports in natural disaster risk I calculate exposure risk as the relative amount of people affected (Birkmann, 2006; UNU-EHS, 2018).

In literature, the various methods used to calculate vulnerability risk encompass quantitative or qualitative data or a combination of both (Birkmann, 2006). A consensus in measuring the vulnerability of a region is yet to be found (Bollin & Hidajat, 2006). In this research natural hazards and the vulnerability of the region must be comparable, therefore I use quantitative data. Even though most papers concerning regional vulnerability risk of natural hazards are focused on developing countries, the Center for Hazards Research (CHR) has developed a method to measure natural risk hazard on census level in the United States. This method is based on earlier research by Odeh and Simpson & Human (2002; 2008) and can also be applied on state level. Simpson & Human (2008) call it the Social Vulnerability Index (SoVI). The SoVI is based on research from Cutter and Morath and the Hazards & Vulnerability Research Institute (HVRI) (2009; 2003). The components of the SoVI changed a little over the years, but the essence remains the same. In 2018 the index is comprised of 15 socioeconomic variables that indicate how a community can reduce the negative impact of a natural disaster. The variables are grouped in four summary themes: Socioeconomic status, household composition & disability, minority status & language and housing type & transportation.

2.2 Segmented Financial Market

This research builds on and adds to a growing body of literature on geographically segmented financial markets in the United States. Bias (1992) for example proves the United States financial market is segmented on a state level. He finds that monetary policies have heterogeneous impacts on economic activity throughout the U.S. In addition to Bias, more regional theorists developed a model of financial markets which prove differences in net wealth have different economic impact in different regions (Harrigan & McGregor, 1987; Moore & Hill, 1982). Becker proves geographic segmentation of the U.S. loan market based on demographic variation in savings (2007). All of these claims suggest the geographical segmentation of the U.S. financial market. Proof of geographic segmentation can also be found in relation to asset pricing. For instance, Pirinsky and Wang (2006) find strong comovement in stock returns of firms headquartered in the same state. They attribute this effect to the geographic component and the segmented U.S. market. Parsons (2016) reports a geographic momentum between co-headquartered firms in the same state across different sectors. This paper also builds on the research on geography-based trading performed by Korniotis and Kumar (2013). When analysing segmented markets, the distinctive regions must be sufficiently heterogeneous in their economic conditions, but enough data has to be available for the region (Bias, 1992). Within the U.S., regions can be distinguished in (from large to small): 9 census divisions, 50 states, 3.142 counties and 74.134 census tracts (Census Bureau, 2010). Like most of the abovementioned studies, a state level analysis is most appropriate for this research. When the selected region becomes smaller, e.g. county or census tract, it becomes impossible to determine the economic activity for a firm, as the data is not available. When the selected regions are census divisions, firms do not have enough variation in economic activity. Most literature on segmented markets look at a firm's headquarter state as their economically relevant region. Comparable to research of Smajlbegovic (2019) I identify all economically relevant states to determine the firm-specific natural hazard risk.

2.3 Relation between Natural Hazard Risk and Returns

I conjecture that natural hazard risk captures a dimension of systematic risk that is not captured by market beta in the Capital Asset Pricing Model (CAPM) or any other of the standard control variables, as the impact of natural disasters to stock returns is mainly local (T. West & G. Lenze, 1994). Previous research relating to natural disasters has shown that the aggregate stock market returns are largely unaffected by big natural disasters (Strobl, 2011; Tavor & Teitler-Regev, 2019; Wang & Kutan, 2013; Worthington, 2008). There is however a body of evidence showing that companies affected by natural disasters find long lasting negative effects to their stock returns (Bourdeau-Brien & Kryzanowski, 2017; T. West & G. Lenze, 1994; Tavor & Teitler-Regev, 2019). Due to this local long lasting impact which is not captured by any of the standard control variables, I hypothesize that investors require compensation for holding stocks with higher natural hazard risk.

2.4 Fama & MacBeth and Portfolio Sorts

For this research, I test the main hypothesis in two separate ways, a Fama and Mcbeth (1973) regression and Portfolio sorts. These two methods are commonly used in asset pricing literature. The Fama and Macbeth (1973) regression method is used to estimate the regression coefficient or beta and the risk premium for different risk factors in asset pricing models.

Perhaps the most well-known asset pricing model in predicting future stock returns is the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966). According to the CAPM excess stock returns of any security are equal to the risk-free rate added with the security's market beta multiplied by the market risk premium. Due to the poor explanatory power of this model I use multiple control variables throughout the analyses. In 1993, Fama and French constructed the three-factor model including the market factor (MKT), the size factor (SML) and the value factor (HML). To improve the power of the three factor model Carhart (1997) added the momentum factor (UMD). In asset pricing literature these four factors are usually augmented with the liquidity factor (LIQ) generated by Pastor and Stambaugh (Pástor & Stambaugh, 2003). These five factors are frequently used in finance literature as risk factors.

Besides controlling for the common control variables, I control the effect of regional natural hazard risk for regional variables that might affect excess returns. For instance, I add a control variable for geographic momentum of Parsons (2016). Furthermore to account for the comovement of firm's returns based on their headquarter location (Pirinsky & Wang, 2006), I divide the firm-specific natural hazard risk in risk within the headquarter state, and risk excluding the headquarter state. In addition, Ellison and Glaeser (1997) show that industries in the U.S. are geographically concentrated. This in combination with Moskowitz and Grinblatt (1999), who show that trading on industry momentum is a profitable strategy, the predictive power of natural hazard risk has to be controlled for industry momentum. As described in section 2.3, the firm-specific natural hazard risk depends on the states in which a company is active. Garcia and Norli (2012) show in their research that truly local companies outperform firms which are geographically dispersed as local firms have lower recognition which yields in higher stock returns. In this research I control for this geographical dispersion by adding two variables to the Fama and MacBeth (1973) regression analysis.

3 Data and Methodology

This chapter provides an overview of the data and methods used to analyze the effect of regional natural hazard risk to stock returns. First I describe the construction of the natural hazard risk proxy, this includes the calculation of the exposure share and the vulnerability score. Second I determine the economically relevant regions for a firm and combine these with natural hazard risk data of the respective states. The next subsection provides the data sources used for the other firm characteristics. Finally, this chapter provides the descriptive statistics.

3.1 Natural Hazard Risk

As explained in the theoretical background section, the natural hazard risk per state is a multiplication of the exposure share and the vulnerability score per state:

$$Natural\ Hazard\ Risk_{s,t} = \left[Exposure\ Share_{s,t}\right] * \left[Vulnerability\ Score_{s,y-1}\right]$$

The natural hazard risk is the predicted risk for state *s* in month *t*. I merge the predicted exposure of month *t* with vulnerability score of the previous year (*y* – 1). This is in correspondence with standard finance procedure and ensures that all the components of natural hazard risk are available to the public at the time *t* when stock returns are predicted.

3.1.1 Exposure Risk

To acquire the data on hazard exposure, I use two databases: the EM-DAT database and the preview Global Risk Data Platform (GRDP) (2019; 2013).

The EM-DAT database has been composed by the Centre of Research on the Epidemiology of Disasters (CRED) and provides core data on 22,000 mass disasters based on various sources, including UN agencies non-governmental organisations, insurance companies, research institutes and press agencies (CRED, 2019). The database gives a good overview on the monthly frequency of natural disasters, but does not always include the exposed population to the natural disasters.

The GRDP has been constructed by Global Resource Information Database Geneva (GRID) and the United Nations Office for Disaster Risk Reduction (UNISDR). The GRDP provides spatial data on the exposed area of natural disasters per square kilometre using a Graphic Information System (GIS). However, it does not give a clear overview on the frequency of the natural disasters. I solve the shortcomings by combining the EM-DAT and the GRDP datasets. The combined dataset consists of spatial maps with the average affected area by natural disasters per month. This automatically adjusts the risk to seasonality; cyclone winds, cyclone surges and floods are more predictable than earthquakes, because their occurrence is seasonal. The combined dataset is constructed using QGIS 3.4 software. I calculated the average affected area and expected frequency based historical natural disaster data. I use historical data of year $y - 20$ to $y - 1$ to predict returns of month t in year y to ensure all the data is available to the public at the time the stock returns are predicted.

I match the combined spatial maps with population data from LandScan. LandScan provides annual population datapoints per square kilometre using GIS. By mapping these datapoints alongside the average affected area, I am able to determine which part of the population is affected by natural disasters based on their location. I use LandScan data of year $y - 1$ to calculate the exposure share in year y to make sure that all the data is available to investors at the time the stock returns are predicted In order to link the datapoints to the relevant state, TIGER/Line Shapefiles produced by the United States Census Bureau have been mapped over the datapoints. The QGIS software allows for a point count analysis in a polygon. By rendering the different U.S. states as polygon and adding the total population to the polygon, I calculate the average affected population per month. I calculate an exposure share for all U.S. states on a monthly basis:

$$\text{Exposure Share} = \frac{n_{p,d,s,t}}{Total\ Pop_{s,y-1}}$$

In the equation $n_{t,p,d,s}$ is the average number of people p affected by the selected natural disasters d in month t in state s, divided by the total population of state s, lagged by one year.

3.1.2 Vulnerability Risk

The vulnerability of the states is based on the SoVI as described in theoretical background. I obtain the social economic variables from the United States Census Bureau and merge the data on state level. Even though the official United States Census is published every ten years, estimates are published on an annual basis. I construct a time series of the vulnerability for each state using the official publications and the annual estimates. The estimates of year y are published in year $y + 1$. For each year I rank all the variables for all the U.S. states on a percentile basis, then I sum all the percentiles for each state and calculate the overall percentile ranking. Percentile ranking values range from 0 to 1, a higher value corresponding with a higher vulnerability risk for the state. Mississippi and New York for instance are among the states that are most vulnerable to a natural disaster. Mississippi is consistently among the most vulnerable states, mostly due to their high poverty and unemployment rate. New York is ranked second since 2000 because, amongst other variables, it is a heavily populated area with a high percentage of minorities. New Hampshire is the state with the lowest vulnerability rank. Most notably this is due to their low poverty rate and the low population density. Table 1 reports the vulnerability ranking for each of the U.S. states in census years.

TABLE 1

Vulnerability Ranking of U.S. States

Table 1 reports the vulnerability scores for each of the U.S states in the census years 1990, 2000 and 2010. The state's vulnerability score is based on 15 census variables. All variables, except income per capita, are ranked from highest to lowest with the highest score corresponding to a higher vulnerability. Income per capita is ranked from highest to lowest as a higher value indicates a lower vulnerability. A percentile rank is calculated for each variable. A percentile rank is the proportion of scores in a distribution that a specific score is greater than or equal to. The percentile ranks are calculated by the following formula $Percentile\ Rank = \frac{Rank-1}{N-1}$. N represents the number of data points. For the final calculation of the vulnerability score, percentile ranks for each state are summed, and an overall percentile rank is calculated by the same formula.

State	1990	2000	2010
Alabama	0.78	0.7	0.7
Alaska	0.82	0.54	0.8
Arizona	0.92	0.88	0.9
Arkansas	0.74	0.72	0.82
California	0.88	0.82	0.96
Colorado	0.18	0.32	0.36
Connecticut	0.26	0.42	0.62
Delaware	0.2	0.5	0.32
District Of Columbia	1	1	1
Florida	0.84	0.94	0.92
Georgia	0.76	0.72	0.88
Hawaii	0.64	0.8	0.22
Idaho	0.32	0.28	0.1
Illinois	0.8	0.76	0.78
Indiana	0.16	0.22	0.44
Iowa	0.04	0.04	0.12
Kansas	0.24	0.2	0.18
Kentucky	0.7	0.46	0.5
Louisiana	0.98	0.84	0.72
Maine	0.22	0.12	0.04
Maryland	0.3	0.52	0.38
Massachusetts	0.48	0.4	0.52
Michigan	0.56	0.46	0.3
Minnesota	0.14	0.14	0.2
Mississippi	0.96	0.96	0.94
Missouri	0.44	0.36	0.56
Montana	0.42	0.24	0.14
Nebraska	0.06	0.1	0.28
Nevada	0.52	0.9	0.76
New Hampshire	0	0.02	0
New Jersey	0.56	0.78	0.6
New Mexico	0.92	0.92	0.84
New York	0.9	0.98	0.98
North Carolina	0.5	0.6	0.68
North Dakota	0.34	0.18	0.16
Ohio	0.34	0.3	0.46
Oklahoma	0.66	0.58	0.64
Oregon	0.4	0.64	0.66
Pennsylvania	0.38	0.38	0.46
Rhode Island	0.54	0.68	0.58
South Carolina	0.68	0.66	0.74
South Dakota	0.6	0.26	0.42
Tennessee	0.62	0.44	0.54
Texas	0.86	0.86	0.86
Utah	0.06	0.06	0.02
Vermont	0.02	0	0.08
Virginia	0.28	0.34	0.24
Washington	0.46	0.62	0.4
West Virginia	0.72	0.56	0.34
Wisconsin	0.1	0.16	0.26
Wyoming	0.12	0.08	0.06

3.2 Regional Economic Relevance

The regional economic relevance of states for a firm have been obtained from a database assigning a weight to economically relevant states based on their citations in annual reports.[1] The data is based on information from the 10-K annual reports found in the Electronic Data Gathering, Analysis and Retrieval (EDGAR) database of the U.S. Securities and Exchange Commission (SEC). The database records the counts for each of the U.S. states in all items of annual reports filed between 1994 and 2014. Based on previous literature I assume regional economic activity to be dependent of the demand for a firms products or services (Nakamura, Steinsson, Barro, & Ursúa, 2013; Smajlbegovic, 2019). Citations linking to production facilities, which are irrelevant for a firm's future cash flow or revenue have been excluded. Based on the citation counts from the annual reports a citation share of the firm's states can be constructed. The citation share is the ratio of citations of a state relative to the citations of the citations of other U.S. states.

$$\text{CitationShare} = \frac{n_{i,s,t}}{\sum_{s=1}^{51} n_{i,s,t}}$$

In the equation $n_{i,s,t}$ is the number of state s counts in the annual report of firm i in year t. The citation share takes a value between 0 and 1 for each firm for each year.

By combining the natural hazard risk of the relevant states with the citation share I calculate the monthly predicted firm-specific natural hazard risk:

$$Natural\ Hazard\ Risk_{i,t} = \sum_{s=1}^{51} CitationShare_{i,s,t-1} \times Natural\ Hazard\ Risk_{s,t}$$

Comparable to other literature I also use the database on citations to construct two variables that explain the cross section of expected stock returns based on geographical location (García & Norli, 2012; Smajlbegovic, 2019). The first of the variables is the state dispersion (STATEDISP). The STATEDISP is the amount of different states named in the annual report of a firm. The second variable is an adaptation of the Hirschmann Herfindahl Index (HHI). The variable sums the squared CitationShares of the different states. A high value of the HHI indicates that the economic relevance for a firm is highly concentrated. In this research logarithm of STATEDISP and HHI are used as control variables.

[1] Database constructed by Esad Smajlbegovic (2019)

3.3 Fama and MacBeth Regression Analysis

In order to examine the cross sectional relation between natural hazard risk and excess stock returns, I employ the Fama and Macbeth (1973) regression analysis. The analysis consists of a two-step procedure. The first step is a cross-sectional regressions of the dependent variable (excess stock returns) on the independent variable (natural hazard risk) and the control variables:

$$RET_{i,t} = \delta_{0,t} + \delta_{1,t}X1_{i,t} + \delta_{2,t}X2_{i,t} + \delta_{3,t} \dots + \cdots + \varepsilon_{i,t}$$

$RET_{i,t}$ is the excess return of stock *i* in month *t. X1, X2*, etc. are the independent variables for month *t.* To prevent that extreme values or skewness of independent variables have a large effect on the dependent variable, some variables are logarithmized. The regression results in the slope coefficients $\delta_{0,t}, \delta_{1,t}$, etc.

The second step of the regression analysis is to calculate the time-series averages of the regression coefficients and examining whether these coefficients are statistically different than zero. Following common asset pricing literature *t* – statistics are adjusted for heteroskedasticity and autocorrelation in the error terms following Newey and West (1987).

3.4 Portfolio Analysis

Besides the Fama and MacBeth regression I perform a portfolio analysis to assess the cross-sectional relationship between natural hazard risk and excess returns. The sort variable is the natural hazard risk. Each month I sort the natural hazard risk variable in five portfolios based on 20% breakpoints. The excess return is calculated both based on equal- (Jegadeesh & Titman, 1993) and value-weighted portfolios (Chui, Titman, & Wei, 2003). To risk adjust the excess returns, I include the Fama and French factors (2013): The excess return of the market portfolio (MKTRF), the size premium (SMB) and the value premium (HML). Besides I include the momentum factor (UMD) and the liquidity factor (LIQ). I obtain these from Kenneth French's website and Lubos Pastor's website respectively.

3.5 Other Stock Characteristics and Data Sources

Stock specific characteristics (monthly stock returns, stock prices, bid-ask spreads, trade volumes and outstanding shares) are obtained from the Center for Research in Security Prices (CRSP). The accounting variables necessary, like book value, the industry and the headquarter location of the firms are retrieved from the CRSP-Compustat merged file (CCM). To match the CCM data with the annual data obtained from CRSP, I use the link-used Table in the

CRSP/Compustat database. The link-used Table matches the Central Index Key (CIK) in the Compustat database to a PERMNO used in CRSP. The final stock sample used in this research are all the common stocks that are listed on the AMEX, NYSE and NASDAQ in the period from July 1995 to June 2014. This average number of firms is around 3,850 per month.

To ensure that the firm characteristics are based on data that would have been publicly available at the time of the analysis, I match the stock specific characteristics from July in year *y* through June in year *y* + 1 to the accounting variables from year *y* – 1. This is according to the standard approach proposed by Fama and French (1993). The same goes for the calculation of the book-to-market ratio (BEME). To ensure that the data is publicly available at the time of analysis, I assume that the BEME that is calculated with data from calendar year *y* is not publicly available until the end of June *y* + 1. For the monthly analysis in this research, the BEME for the months *t* from July *y* + 1 till June *y* + 2 is taken from the book value of equity (BE) measured at the end of the fiscal year ending in calendar year *y* is divided by the market value of equity (ME) at the end of December of calendar year *y*.

3.6 Summary Statistics

Table 2 presents the descriptive statistics for the state related variables, other firm characteristics and standard asset pricing models used in this study. The natural hazard risk variables in Panel A (NHRISK, EXPOSURE & VULNERABILITY) are equally weighted based on monthly expectations calculated with equations 1 to 3. The sample firms are on average active in 11 different U.S. states, with 9 states as median. The mean and median HHI amount to 0.356 and 0.305 respectively. The return data from Panel B is monthly return data based on the closing prices at the end of the month. The book value (BE) is calculated on an annual basis.

TABLE 2

Summary Statistics of Explanatory Variables and Risk Factors

Table 2 reports the summary statistics of state-related variables, other firm characteristics and standard asset pricing factors in the period July 1995 to June 2014. The reported statistics are the mean, standard deviation, 1st, 25th, 50th, 75th, and 99th percentile, and the number of available observations. Panel A. reports the variables which are state-related: The natural hazard risk (NHRISK), the exposure share (EXPOSURE), the vulnerability score (VULNERABILITY) to natural hazards, state dispersion (STATEDISP) represents the number of distinct states which are cited in a firm's annual reports, and the Herfindahl-Hirschman concentration measure based on state citations (HHI). Panel B displays other firm characteristics. This includes market beta (βMKTRF) and the βSMB and βHML. The betas are calculated with daily data using rolling regressions for the past 125 days. ISVOLA is the standard deviation of the regressions residuals (Ang et al. (2009)). In addition to the standard market capitalization (MKTCP) calculated each month, Panel B reports MKTCPff,cpi: the share price times the number of shares outstanding calculated as of the end of the most recent June adjusted using the consumer price index to reflect June 2014 dollars (Fama and French (1992)). In a similar manner the book value of common equity (BE) and the book-to-market ratio (BEME) are calculated in accordance with Fama and French (1992). The bid-ask spread (BIDASK) is computed using daily data over a six month window as in Amihud and Mendelson (1986). RET_1 and RET2_12 display respectively the one-month lagged excess return (Jegadeesh (1990)) and the cumulative excess return from month $t-12$ to $t-2$ (Jegadeesh and Titman (1993)). INDRET_1 and INDRET2_12 display the lagged excess returns for the firm's industry using the 49 Fama-French industry classification (Moskowitz and Grinblatt (1999)). HQRET_1 denotes the lagged average excess return for all companies with a headquarter in the same state as the specific firm (equal weighted) (Parsons et al (2017)). Furthermore, Panel C displays the standard asset pricing statistics MKTRF, SMB, HML UMD and LIQ (Fama and French (1993), Carhart (1997), and Pástor and Stambaugh (2003)).

			percentile					
Variable	Mean	St.Dev	p1	p25	Median	p75	p99	n
Panel A. State-Related Variables								
NHRISK	0.020	0.026	0.001	0.004	0.009	0.023	0.111	808616
EXPOSURE	.024	.029	.002	.007	.013	.027	.120	808616
VULNERABILITY	.624	.164	.198	.509	.637	.747	.929	808616
STATEDISP	11.311	8.763	2.000	6.000	9.000	14.000	47.000	808616
HHI	0.356	0.209	0.060	0.200	0.305	0.463	0.941	808616
Panel B. Other Firm Characteristics								
βMKTRF	0.836	0.888	-1.416	0.349	0.831	1.280	3.232	801724
βSMB	0.663	1.181	-2.141	0.007	0.558	1.227	4.046	801724
βHML	0.256	1.525	-3.943	-0.402	0.234	0.930	4.413	801724
MKTCPff,cpi	3611.302	18781.708	5.589	73.525	298.847	1352.169	65017.516	807976
MKTCP	3052.292	15768.442	3.963	57.251	236.710	1103.062	55134.055	802392
BE	1318.776	7262.829	-105.026	31.050	122.977	528.184	21084.031	806315
BEME	0.771	1.035	0.043	0.322	0.562	0.914	4.258	740250
ISVOLA	0.033	0.026	0.007	0.017	0.026	0.041	0.125	802616
BIDASK	0.023	0.037	0.000	0.002	0.010	0.028	0.167	798369
RET_1	0.011	0.187	-0.410	-0.071	0.000	0.073	0.608	800095
RET2_12	0.124	0.793	-0.844	-0.238	0.017	0.298	2.910	807929
INDRET_1	0.011	0.079	-0.207	-0.029	0.012	0.049	0.243	808616
INDRET2_12	0.136	0.343	-0.505	-0.069	0.100	0.276	1.370	807929
HQRET_1	0.011	0.069	-0.176	-0.028	0.013	0.050	0.207	807927
Panel C. Standard Asset Pricing Factors								
MKTRF	0.476	4.703	-10.720	-2.360	1.190	3.610	9.540	808616
SMB	0.332	3.624	-6.980	-1.550	0.130	2.540	10.610	808616
HML	0.336	3.653	-10.180	-1.450	0.280	2.200	12.290	808616
UMD	0.544	5.740	-16.170	-1.350	0.770	3.200	16.590	808616
LIQ	0.794	3.859	-9.705	-1.271	0.751	3.054	10.203	808616

In order to show a comprehensible relation of stock returns and other firm characteristics from Table 2 to NHRISK, I sort NHRISK in quintiles. For each month I divide the cross section in quintiles depending on the value of NHRISK. Quintile 1 corresponds with the lowest values, and quintile 5 with the highest values of NHRISK. For the other variables I compute the average value across time within the quintile. The results of the quintile sort are shown in Table 3. Due to the simplicity of the sorting method, any implied relation between the variables should be subjected to more thorough testing before conclusions can be made. I examine the dependencies of the variables further in section 4 of this research.

The quintile sort suggests a positive relation between NHRISK and monthly excess returns. But it also shows that the market beta and SML beta are increasing across the natural hazard risks quintiles. This is comparable to the lagged returns and industry returns. The HML beta and the bid ask spread appear to be decreasing as the quintiles increase. The market capitalization and the state dispersion seem to be at their lowest in the first quintile. The highest values are found in the third quintile, after which the market capitalization and the state dispersion decline again. The HHI shows that firms with highest level of regional concentration are located in the first and the fifth quintile, with lower values in quintile 3 and 4. In short companies in the lowest quintile are on average the smallest, regionally concentrated companies with relatively low exposure to the market portfolio. Companies in the highest natural risk quintile are also small and regionally concentrated, but they are relatively exposed to the market portfolio and the SMB beta.

TABLE 3

Natural Hazard Risk and Other Explanatory Variables

Table 3 displays the mean value of variables within each quintile of natural hazard risk (NHRISK). For each month, the cross section is divided in quintiles depending on the value of NHRISK. Variables are described in Table 2. The sample period is July 1995 to June 2014.

	Natural Hazard Risk Quintiles				
Variable	1	2	3	4	5
NHRISK	0.003	0.005	0.009	0.019	0.063
RET	0.009	0.010	0.010	0.011	0.012
STATEDISP	8.262	11.520	13.985	13.738	9.052
HHI	0.459	0.358	0.306	0.277	0.379
βMKTRF	0.679	0.802	0.863	0.894	0.942
βSMB	0.567	0.615	0.634	0.672	0.828
βHML	0.307	0.322	0.317	0.266	0.067
MKTCP	1466.962	2709.655	4395.837	4057.067	2635.807
MKTCPff,cpi	1785.028	3240.785	5269.304	4719.815	3042.283
ISVOLA	0.030	0.032	0.032	0.033	0.038
BIDASK	0.025	0.024	0.022	0.021	0.021
RET_1	0.009	0.010	0.011	0.011	0.012
RET2_12	0.104	0.122	0.119	0.128	0.146
INDRET_1	0.009	0.010	0.011	0.011	0.011
INDRET2_12	0.123	0.133	0.134	0.139	0.150
HQRET_1	0.010	0.010	0.010	0.010	0.012

4 Natural Hazard Risk and Stock Returns

In this chapter I analyze the effect of firm-specific natural hazard risk to stock returns using two well-known methods in finance literature: Fama and MacBeth (1973) regressions and portfolio sorts. In the next subsection I perform robustness tests to examine the stability of the relation between natural hazard risk and stock returns.

4.1 Regression-Based Tests

First I conduct the Fama and French (1993) regression following the Fama and MacBeth (1973) regression analysis. Monthly excess return is the dependent variable and the natural logarithm of NHRISK is the independent natural hazard risk variable. Control variables are added depending on the specification. The time-series average of each regression coefficient and the corresponding *t* – statistic are calculated using the Newey and West (1987) standard error correction.

In the first column of Table 4, I perform the regression with the natural logarithm of natural hazard risk [ln(NHRISK)] as only explanatory variable. The calculated regression coefficient of 0.001 is slightly positive, but does not significantly predict excess returns as the *t* value is 0.80.

The second column includes the standard control variables to the Fama and MacBeth (1973) framework. When controlling for the standard control variables I find a similar regression coefficient of ln(NHRISK) of 0.001. However, the control variables increase the statistical significance to 5% with a *t* – statistic of 2.01. To interpret this result economically I use the sorting exercise from Table 3. The average difference in the coefficient between the lowest and the highest quintile of natural hazard risk yields an increase of average excess stock return of .304 percentage points. The regression coefficients of the controlling variables are as expected from previous asset pricing literature. The historical market beta, the bid-ask spread, cumulative previous returns and idiosyncratic volatility do not have a significant coefficient. However, firm size and the book-to-market ratio show a strong effect in the expected direction. The short-term reversal displays the most substantial coefficient.

In the third column I include control variables for the industry momentum (Moskowitz & Grinblatt, 1999) and the lead-lag effect of returns in the headquarter state, or geographic momentum (Parsons, 2016). Control variables for industry momentum are industry return for the past month and cumulative past industry return for the previous year. The control variable for geographic momentum, is the past month's return for firms headquartered in the same state.

Even though the three variables significantly influence the excess returns, the natural hazard risk coefficient remains significant at the 5% level without changing the coefficient.

In order to account for research showing local firms outperforming firms that are active in more U.S. states (García & Norli, 2012), I include state dispersion and Herfindahl-Hirschmann Index variables. These are tested in column 4 and 5 of Table 3. In my research I do not find significant coefficients for either the natural logarithm of STATEDISP or HHI. This result could be influenced by the different sample period. The coefficient of the natural logarithm of natural hazard risk remains 0.001 with a significance level of 5%. This shows that the predictive power of natural hazard risk is not due to state dispersion.

In column 6 of Table 4 I control for comovement in excess returns of firms headquartered in the same state. I divide the natural logarithm of natural hazard risk in two variables, the natural hazard risk in the headquarter state and the natural hazard risk in all relevant states excluding the headquarter state. The results show that the coefficient of both proxies is 0.000, moreover the t – statistics of 0.93 and 0.72 respectively are not significant by standard confidence levels. These results implicate that the natural hazard risk proxy only explains the excess returns when all states are included. The return comovement of firms with headquarters in the same state do not explain the predictive power of natural hazard risk.

TABLE 4

Natural Hazard Risk and Stock Returns

Table 4 presents the results of Fama-MacBeth (1973) regressions of monthly stock returns on natural hazard risk [ln(NHRISK)]. The dependent variable is the excess stock return in month *t*, denominated in U.S. dollars. For each stock the predicted regional natural hazard risk in the headquarter state is calculated [ln(NHRISKHQ)], and the predicted natural hazard risk for all relevant states except the headquarter state [ln(NHRISKEXHQ)]. Other firm characteristics are described in Table 1. . Reported coefficients are time-series averages of the cross-sectional regressions shown above, with t-statistics calculated based on the Newey-West (1987) standard errors of these estimates. Stars are used to show significance level: **** p<0.01, ** p<0.05, * p<0.1*. The sample period is July 1995 to June 2014.

Variable	1	2	3	4	5	6
ln(NHRISK)	0.001	0.001**	0.001**	0.001*	0.001**	
	(0.80)	(2.01)	(2.02)	(1.97)	(1.98)	
ln(NHRISKHQ)						0.000
						(0.93)
ln(NHRISKEXHQ)						0.000
						(0.72)
βMKTRF		0.002	0.002	0.002	0.002	0.002
		(0.85)	(0.89)	(0.88)	(0.86)	(0.91)
ln(SIZE)		-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
		(-3.09)	(-3.10)	(-3.14)	(-3.10)	(-3.10)
ln(BEME)		0.002*	0.002**	0.002*	0.002*	0.002**
		(1.89)	(2.21)	(1.92)	(1.92)	(2.16)
ln(ISVOLA)		0.002	0.001	0.001	0.002	0.001
		(0.47)	(0.38)	(0.43)	(0.45)	(0.45)
RET_1		-0.038***	-0.044***	-0.038***	-0.038***	-0.044***
		(-7.19)	(-8.69)	(-7.32)	(-7.27)	(-8.80)
RET2_12		0.000	-0.001	0.000	0.000	-0.001
		(-0.06)	(-0.43)	(-0.05)	(-0.06)	(-0.42)
ln(BIDASK)		0.000	0.000	0.000	0.000	0.000
		(-0.12)	(0.20)	(0.06)	(0.11)	(0.24)
INDRET_1			0.138***			0.134***
			(8.13)			(8.06)
INDRET2_12			0.013***			0.012**
			(2.64)			(2.55)
HQRET_1			0.039***			0.044***
			(3.10)			(3.32)
ln(STATEDISP)				0.000		0.000
				(0.16)		(0.27)
HHI					-0.001	-0.001
					(-0.36)	(-0.19)
Constant	0.014*	0.026	0.02	0.026	0.026	0.019
	(1.88)	(1.61)	(1.42)	(1.56)	(1.62)	(1.22)
Observations	800865	727315	727315	727315	727315	717606
R-squared	0.005	0.067	0.074	0.068	0.068	0.077

4.2 Portfolio Tests

In this subsection I examine the profitability of a long and short portfolio in natural hazard risk over the entire sample period. Table 5 reports the average excess return portfolios based on natural hazard risk quintiles. The first five columns shows excess returns of portfolios from the lowest natural hazard risk (column 1) and the highest natural hazard risk (column 5). Column 6 shows the excess returns of a strategy that goes long in the highest risk quintile and short in the lowest risk quintile. Panel A displays the results of an equal weighted portfolio and Panel B reports the value weighted portfolio. Both Panels show a positive coefficient in the sixth column indicating a monthly return of 0.215% and 0.331% respectively. The t – statistics however show no significance by any standard confidence levels.

TABLE 5

Portfolio Sorts Based on Natural Hazard Risk

Table 5 presents monthly excess returns for portfolios sorted according to quintiles of natural hazard risk. The sixth column (High – Low) displays the excess returns of a portfolio that goes long in the high risk portfolio and short in the low risk portfolio. Panel A reports equal-weighted excess returns, and Panel B the value-weighted excess returns. The *t*-statistics are in parenthesis. The sample period is July 1995 to June 2014.

	Low NHRISK	2	3	4	High NHRISK	High – Low
Panel A. Equal Weighted						
RETRF	0.986	1.079	1.116	1.127	1.201	0.215
	(3.06)	(2.88)	(2.83)	(2.65)	(2.31)	(0.81)
Panel B. Value Weighted						
RETRF	0.590	0.617	0.551	0.690	0.750	0.160
	(2.08)	(1.92)	(1.88)	(2.35)	(1.62)	(0.47)

In order to adjust the results from Table 5 for to risk, I run a time-series regression with the five well known risk variables in Table 6. Similar to Table 5, Panel A reports equal-weighted portfolios and Panel B reports value-weighted portfolios. The columns of Table 6 are structured as follows: per three columns, the first column shows the low risk portfolio, the second column reports the high risk portfolio and the last column reports the coefficients of a portfolio that goes long in the high risk portfolio and short in de low risk portfolio. In the first three columns I control for the market risk (Fama & French, 1993), column 4 – 6 include SMB and HML factors (Fama & French, 1993), column 7 – 9 add the momentum factor (Carhart, 1997), and finally column 10 – 12 complete the five factor model with the liquidity factor (Pástor & Stambaugh, 2003).

Notably the alphas of the equal- and value-weighted long and short portfolio when controlling for market risk are negative (-0.107; -0.139) with insignificant *t* – statistics of -0.49 and -0.45 respectively. All other α long short portfolios show a positive coefficient. Even though the positive coefficients of α suggest a positive relation between natural hazard risk and excess returns, the *t* – statistics corresponding to the alphas are insignificant by standard confidence levels. Even when controlling for market risk, Fama and French factors and the momentum factor, the *t* – statistics corresponding to the long short portfolio's α coefficient are not significant by standard confidence levels.

TABLE 6

Time-Series Regression of Natural Hazard Risk

Table 6 reports the coefficient estimates of time-series regressions adjusted with five well known asset pricing factors: MKTRF , SMB, HML, UMD and LIQ. The variables are described in Table 1. The α represents Jensen's alpha. The columns with title 'Low' report the coefficient estimates for the lowest natural hazard risk quintile, the columns with title 'High' report the coefficient estimates corresponding to the highest natural hazard risk quintile. The 'High – Low' columns display the excess returns of a portfolio that goes long in the high risk portfolio and short in the low risk portfolio. *t* - Statistics are reported in parentheses. The sample period is July 1995 to June 2014.

Variable	Low 1	High 2	High - Low 3	Low 4	High 5	High - Low 6	Low 7	High 8	High - Low 9	Low 10	High 11	High - Low 12
Panel A. Equal-Weighted Portfolio												
α	0.428	0.320	-0.107	0.219	0.295	0.076	0.329	0.510	0.182	0.323	0.504	0.181
	(2.43)	(1.08)	(-0.49)	(2.08)	(1.62)	(0.56)	(3.47)	(3.23)	(1.39)	(3.38)	(3.16)	(1.38)
MKTRF	0.891	1.406	0.515	0.839	1.188	0.350	0.780	1.073	0.293	0.778	1.071	0.293
	(23.46)	(21.88)	(10.97)	(35.48)	(28.99)	(11.40)	(34.96)	(28.87)	(9.56)	(34.01)	(28.09)	(9.32)
SMB				0.631	0.965	0.334	0.650	1.003	0.353	0.650	1.002	0.353
				(19.51)	(17.21)	(7.96)	(22.53)	(20.86)	(8.89)	(22.47)	(20.80)	(8.87)
HML				0.295	-0.356	-0.651	0.245	-0.455	-0.700	0.245	-0.455	-0.700
				(9.11)	(-6.32)	(-15.46)	(8.28)	(-9.24)	(-17.22)	(8.28)	(-9.21)	(-17.18)
UMD							-0.143	-0.281	-0.138	-0.144	-0.282	-0.138
							(-7.78)	(-9.17)	(-5.46)	(-7.78)	(-9.15)	(-5.43)
LIQ										0.013	0.015	0.002
										(0.55)	(0.37)	(0.05)
Panel B. Value-Weighted Portfolio												
α	0.077	-0.062	-0.139	-0.044	0.149	0.193	-0.057	0.181	0.238	-0.075	0.172	0.247
	(0.62)	(-0.27)	(-0.45)	(-0.43)	(0.93)	(0.92)	(-0.55)	(1.12)	(1.13)	(-0.72)	(1.05)	(1.16)
MKTRF	0.837	1.324	0.487	0.894	1.174	0.280	0.901	1.157	0.255	0.893	1.153	0.259
	(30.97)	(27.10)	(7.26)	(38.63)	(32.76)	(5.98)	(36.62)	(30.44)	(5.15)	(35.60)	(29.61)	(5.11)
SMB				-0.009	0.236	0.245	-0.011	0.242	0.253	-0.012	0.241	0.254
				(-0.28)	(4.82)	(3.83)	(-0.35)	(4.93)	(3.95)	(-0.39)	(4.91)	(3.95)
HML				0.330	-0.665	-0.995	0.336	-0.679	-1.016	0.337	-0.679	-1.016
				(10.40)	(-13.51)	(-15.50)	(10.33)	(-13.50)	(-15.48)	(10.37)	(-13.48)	(-15.45)
UMD							0.017	-0.042	-0.059	0.014	-0.044	-0.058
							(0.85)	(-1.35)	(-1.46)	(0.71)	(-1.39)	(-1.42)
LIQ										0.041	0.021	-0.020
										(1.51)	(0.50)	(-0.37)

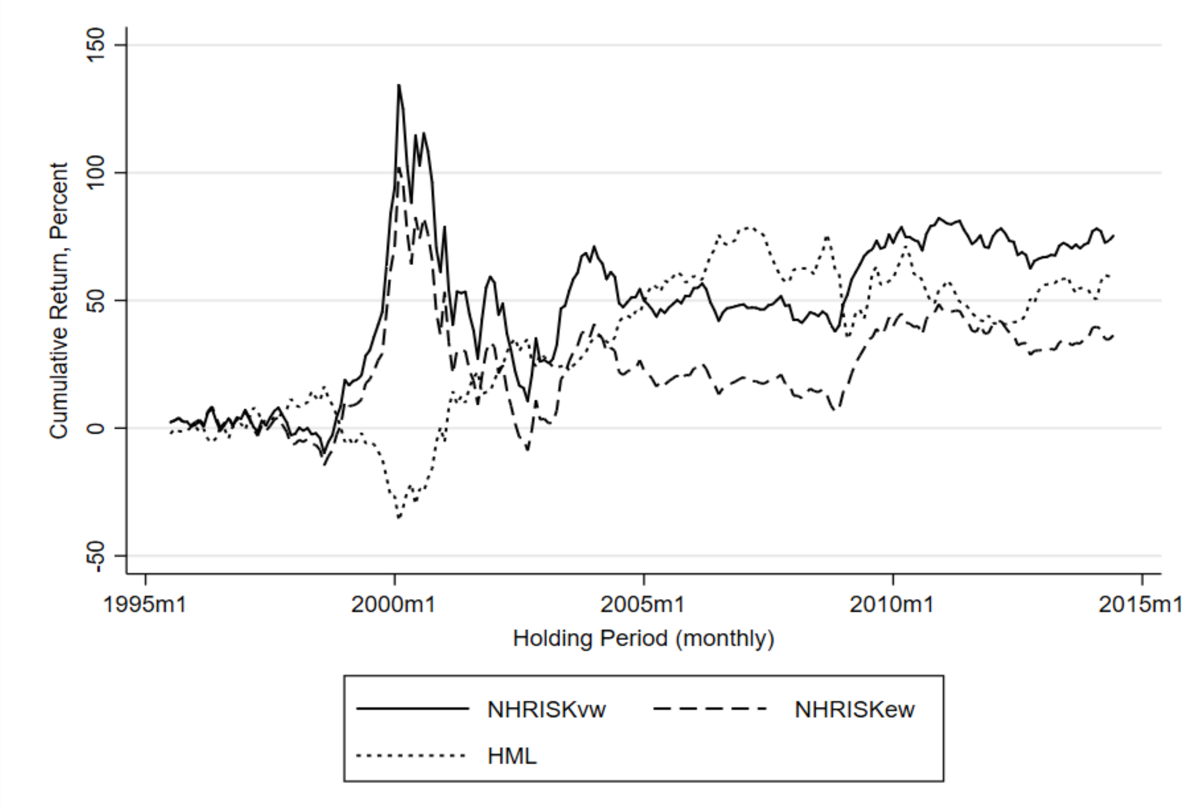
When looking at Table 6 in more detail, it becomes clear that high natural hazard risk portfolios have a higher coefficient to excess market returns than low risk portfolios. Indicating that stocks with a higher natural hazard risk, also have a high systematic risk. This could partly explain the excess returns found in Table 5. Besides the exposure to the market, the long short portfolios are positively related to the SMB factor. The positive returns of the SMB factor means that the high risk portfolio holds more weight to smaller capitalization stocks. It stands out that the HML coefficient on the long - short portfolio is negative. For instance, in Panel A and B the HML coefficient for the five-factor model is -0.700 and -1.016 respectively with highly significant *t* – statistics of -17.18 and -15.45. In general, the HML coefficient is positive when the portfolio is relatively more exposed to value stocks, and negative when the factor is more exposed to growth stocks. The negative coefficient of HML appears to be driven by purely by negative coefficient of the high risk portfolios, as the low risk portfolio's coefficient to HML is positive. This indicates that a portfolio formed with high natural hazard holds more weight to growth stocks. The long short portfolios are negatively related to the momentum factor, even though this relation is not significant at the value-weighted portfolio. The long short portfolios has no significant tilt toward the liquidity factor.

Based on the indication that an investment strategy based on natural hazard risk might be a hedge to an HML investment strategy, I compare the returns from the equal- and value weighted NHRISK long short portfolios to the HML portfolio in Figure 1. Figure 1 shows that the performance of the portfolios in the sample period is very similar. However, the NHRISK long short portfolio moves in opposite direction as the HML portfolio in positive or negative excess return peaks. For instance, while the NHRISK portfolios peak in excess returns, the HML portfolio experienced a severe crash leading up to 2000.

FIGURE 1

Cumulative Return of Natural Hazard Risk and the HML portfolio

Figure 1 presents the cumulative performance of the equal- and value-weighted (NHRISKew, NHRISKvw) long and short portfolio in NHRISK and the HML portfolio. The sample period is July 1995 to June 2014.



5 Robustness Tests and Additional Insights

In this section, I assess the structural validity of the Fama and Macbeth (1973) regression analysis and the portfolio sorts by performing multiple robustness tests. In addition to the robustness tests I include tests that give additional insights in the effect of natural hazard risk to excess stock returns and I discuss possible alternate explanations of the results.

First I assess the results from the Fama and MacBeth (1973) regression analysis of section 4.1. To eliminate the possibility that the results originate from the sample selection bias, I use different specifications for the sample selection and look at alternative sample periods while controlling for the standard controls as in column 2 of Table 4. As financial industry stocks are have a highly different business model than other companies, they might influence the Fama and MacBeth (1973) regressions (Fama & French, 1993). To ensure that financial industry stocks don’t drive the results, it is common in asset-pricing literature to exclude them from the sample. I present the results in column 1 of Table 7. The coefficient of the natural logarithm of natural hazard risk remains the same at 0.001. The *t* – statistic is to 1.70 which is statistically significant at the 10% level, even though the explanatory power is slightly lower than with the financial sector included in the sample (2.01).

As the portfolio sorts in Table 6 show, stocks with higher natural hazard risk tend to be negatively exposed to the momentum factor (UMD). Bhootra (2011) shows that the momentum portfolio significantly increases when penny stocks trading below $5 per share are excluded from the sample. The research of Bhootra (2011) demonstrates that failure to exclude penny stocks could result in a downward bias of the momentum factor. To make sure this bias does not influence the results, I exclude penny stocks or microcaps from the sample. Column 2 reports on coefficients without shares trading at a price of less than $1, and in column 3 shares with a share price below $5 are excluded. The statistical and economical significance of natural hazard risk does not change without microcaps while controlling for the standard control variables.

Columns 4 – 6 of Table 7 resent the results of the regression analysis for three separate time periods. These time periods are July 1995 to August 2008, September 2008 to January 2011 and February 2011 to June 2014 in columns 4, 5 and 6 respectively. The middle time period (column 5) corresponds to market decline as a consequence of the global financial crisis in 2008. As described above, the natural hazard risk portfolio might behave as an hedge to the HML portfolio. According to Jegadeesh and Titman (2016) the negative serial correlation in the period following the global financial crisis led to negative results for the HML portfolio.

The negative exposure to the HML portfolio would lead to expect that the coefficient of ln(NHRISK) increases during this time period. As expected, the coefficient and the statistical significance increase during the crisis years. It is interesting to see the difference in the time periods before and after the financial crisis of 2008. The time period between 1995 and 2008 yields a similar coefficient as column 2 of Table 4, even though the *t* – statistic has declined, after the crisis the effect of natural hazard risk seems to disappear.

Table 7

Fama and MacBeth (1973) Regression Analysis

Table 7 presents the results from the Fama and MacBeth (1973) regression analysis with standard control variables and different sample selections. In the first column the coefficient from the natural logarithm of natural hazard risk is given based on the same sample from Table 4, but excluding financial companies. In the second and third column the penny stocks are excluded from the sample, respectively firms with a stock price under one dollar and five dollars The sample period for the first three columns is July 1995 to June 2014. The sample period for column 4 is July 1995 to August 2008. The sample period for column 5 is September 2008 to January 2011. The sample period for column 6 is February 2011 to June 2014. The *t* – statistics are reported in parentheses.

Variable	Without Financial Sector 1	Stock price >1$ 2	Stock price >5$ 3	1995-2008 4	2008-2011 5	2011-2014 6
ln(NHRISK)	0.001	0.001	0.001	0.001	0.002	0.000
	(1.70)	(1.69)	(1.70)	(1.58)	(2.01)	(0.17)
constant	0.029	0.014	-0.000	0.026	0.030	0.018
	(1.71)	(0.92)	(-0.00)	(1.29)	(0.59)	(0.77)
Standard Controls	YES	YES	YES	YES	YES	YES
Obs.	581909	677840	547411	554754	80700	97392
R-squared	0.064	0.070	0.082	0.071	0.075	0.049

In order to determine the robustness of the portfolio time-series regression from Table 6, I add asset pricing factors and I use alternative breakpoints in assigning the portfolios.

In columns 1 – 3 of Table 8, I add the profitability (RMW) and the investment (CMA) factors (Fama & French, 2015) to the time-series regression. Due to the negative exposure of the natural hazard risk portfolio to the profitability factor, the α in the equally-weighted as well as the value-weighted portfolio are larger than reported in Table 6. Besides, there is a significant increase in the corresponding *t* – statistic.

In columns 4-6 I adjust the breakpoints of the different portfolios. In the first portfolio sorts (Table 5 and 6) the portfolios are based on different natural hazard risk quintiles, or a breakpoint of 20%. To test the robustness I set the breakpoint at 10%, which matches a decile portfolio analysis. The results are comparable with a slightly increased alpha for both the equal-weighted and the value-weighted portfolio.

Table 8

Portfolio Time-Series Regression

Table 8 reports the coefficient estimates of time-series regressions similar to Table 6. In the first three columns the asset pricing factors RMW and CMA have been included. Columns 4-6 reports the coefficient estimates of time-series regressions of portfolios based on natural hazard risk deciles (as opposed to quintiles). The sample period is July 1995 to June 2014.

	Low	High	High – Low	Low (dec)	High (dec)	High – Low (dec)
Variable	1	2	3	4	5	6
Panel A. Equal-Weighted Portfolio						
α	0.323	0.683	0.360	0.345	0.619	0.275
	(3.22)	(4.48)	(2.98)	(3.37)	(3.47)	(1.69)
MKTRF	0.777	0.997	0.219	0.703	1.054	0.351
	(30.39)	(25.60)	(7.12)	(28.68)	(24.69)	(9.00)
SMB	0.646	0.84	0.195	0.607	1.055	0.448
	(19.46)	(16.64)	(4.87)	(19.58)	(19.54)	(9.08)
HML	0.242	-0.329	-0.571	0.283	-0.497	-0.78
	(5.34)	(-4.75)	(-10.43)	(8.91)	(-8.99)	(-15.44)
UMD	-0.144	-0.268	-0.124	-0.137	-0.242	-0.105
	(-7.62)	(-9.33)	(-5.46)	(-6.90)	(-7.02)	(-3.35)
LIQ	0.014	0.032	0.018	-0.007	0.014	0.021
	(0.57)	(0.85)	(0.60)	(-0.28)	(0.30)	(0.51)
RMW	-0.009	-0.460	-0.451			
	(-0.18)	(-6.22)	(-7.71)			
CMA	0.013	0.065	0.052			
	(0.21)	(0.70)	(0.72)			
Panel B. Value-Weighted Portfolio						
α	-0.208	0.369	0.577	0.084	0.470	0.554
	(-1.97)	(2.28)	(2.78)	(-0.57)	(2.01)	(1.87)
MKTRF	0.944	1.073	0.128	0.816	1.247	0.431
	(35.11)	(25.96)	(2.42)	(23.02)	(22.27)	(6.08)
SMB	0.022	0.120	0.098	0.018	0.331	0.314
	(0.64)	(2.24)	(1.42)	(0.39)	(4.68)	(3.50)
HML	0.177	-0.492	-0.669	0.412	-0.834	-1.246
	(3.70)	(-6.70)	(-7.10)	(8.96)	(-11.51)	(-13.57)
UMD	-0.002	-0.024	-0.022	0.058	-0.026	-0.084
	(-0.12)	(-0.79)	(-0.56)	(2.03)	(-0.57)	(-1.46)
LIQ	0.046	0.027	-0.02	0.023	0.033	0.01
	(1.77)	(0.66)	(-0.38)	(0.61)	(0.55)	(0.13)
RMW	0.165	-0.390	-0.555			
	(3.23)	(-4.96)	(-5.50)			
CMA	0.239	-0.128	-0.366			
	(3.76)	(-1.31)	(-2.92)			

To gain further understanding of the effect of natural hazard risk I perform some additional tests. First, I study whether the found effects of natural hazard risk could be attributed exclusively to either exposure to natural hazards or to vulnerability. Because the exposure to natural hazards is rightly skewed, I logarithmize the variable. Table 9 shows positive coefficients for ln(EXPOSURE) and VULNERABILITY to the excess returns of stock prices. The *t* – statistics for the variables are 1.83 and 1.18 respectively, which indicates that the variables by itself do not significantly impact the excess return.

I expect that more local firms have a higher natural hazard risk as they are more exposed to single states. Even though I find that natural hazard risk is more present at local firms (Appendix C), I can find no evidence that this drives any of the results. I test this by using the geographical variables of state dispersion and HHI and test them in an interaction with the logarithm of natural hazard risk within the Fama and MacBeth (1973) framework.

Table 9

Additional Fama and MacBeth (1973) regressions

Table 9 reports the coefficient estimates of Fama and MacBeth (1973) regressions while controlling for the standard controls (column 2 Table 4). The first two columns report the coefficient for the interaction terms between the natural logarithm of NHRISK and the geographic control variables STATEDISP and HHI respectively. Column 3 and 4 report regression coefficients of the natural logarithm of EXPOSURE and the coefficient of VULNERABILITY respectively as independent variables excluding ln(NHRISK). The sample period is July 1995 to June 2014.

	State Dispersion	HHI	Exposure	Vulnerability
Variable	1	2	3	4
ln(NHRISK)	0.001	0.001		
	(2.02)	(1.60)		
ln(EXPOSURE)			0.001	
			(1.84)	
VULNERABILITY				0.002
				(1.18)
STATEDISP x ln(NHRISK)	-0.000			
	(-0.11)			
HHI x ln(NHRISK)		0.000		
		(0.31)		
constant	0.026	0.026	0.027	0.021
	(1.61)	(1.63)	(1.65)	(1.53)
Standard Controls	YES	YES	YES	YES
Obs.	727315	727315	706018	706018
R-squared	0.068	0.068	0.068	0.067

Even though this research suggests that natural hazard risk is a factor in explaining the cross section of excess returns, the explanatory power of natural hazard risk does not exceed the proposed t – statistic of 3.0 (Harvey et al., 2016).. Therefore natural hazard risk cannot irrefutably be considered a factor in explaining cross-sectional results.

The lack of explanatory power of natural hazard risk could be explained by the method of calculating regional natural hazard risk. In this research I use predictions based on historical data. However, in recent meteorological research, multiple empirical prediction algorithms are being developed to more precisely perform multi-season forecasts of North Atlantic hurricane activity (Bender et al., 2010; Caron, Jones, & Doblas-Reyes, 2014; Schumacher & Strobl, 2011). It would be interesting to see whether the predictability of the model increases when the risk of natural hazards can be estimated more accurately.

It could also be possible that the found relation between natural hazard risk and excess return actually rests on mispricing driven by irrational investors and demand shocks based on a (unknown) regional information source. Contrary to the Efficient Market Hypothesis, researchers have found that psychological factors can create anomalies in expected excess returns (Lamont & Thaler, 2003; Shleifer & Vishny, 1997). If the effect of natural hazard risk on excess returns is higher amongst difficult to arbitrage stocks, this could be an indication that the effect found in this research might actually depend on mispricing (Gromb & Vayanos, 2010; Zhang, 2007).

To explore whether the effect might be sustained by difficulties to arbitrage, in Table 10 I use variables relating to costly arbitrage in interaction with the logarithm of natural hazard risk within the Fama and MacBeth (1973) framework. The variables I use are the ln(BIDASK), ln(SIZE) and the ln(ISVOLA). These variables are well-known for determining difficulty to arbitrage. To make the variables easier to interpret, I standardize them with a mean of 0 and a standard deviation of 1. The coefficients in Table 10 corresponding to the interaction terms implicate that limits to arbitrage could play a role in the return predictability of natural hazard risk. Even though the interaction between the logarithmized bid-ask spread and the logarithm of natural hazard risk is not significant by any of the standard confidence levels, the $t-$statistics for size and idiosyncratic risk are significant at a 10% level.

Table 10

Additional Fama and MacBeth (1973) regressions

Table 10 reports the coefficient estimates of Fama and MacBeth (1973) regressions while controlling for the standard controls (column 2 Table 4). The variables used for the interaction with ln(NHRISK) are standardized with a mean of zero and a standard deviation of 1. The sample period is July 1995 to June 2014.

Variable	1	2	3
ln(NHRISK)	0.001	0.001	0.001
	(1.94)	(1.73)	(1.83)
$\ln(BIDASK)_{STD}$ X ln(NHRISK)	0.001		
	(1.54)		
$\ln(SIZE)_{STD}$ X ln(NHRISK)		-0.001	
		(-1.83)	
$\ln(ISVOLA)_{STD}$ X ln(NHRISK)			0.001
			(1.83)
constant	0.035	0.032	0.044
	(1.94)	(1.80)	(1.80)
Standard Controls	YES	YES	YES
Obs.	727315	727315	727315
R-squared	0.068	0.068	0.068

The conjecture of regional mispricing has already been studied within the United States. For instance, Korniotis and Kumar (2013) find abnormal performance in geography-based trading between U.S. state portfolios based on mispricing. When U.S. states have higher unemployment rates and housing collateral ratios are lower, expected returns are higher. They attribute the effects to local risk aversion and arbitraging patterns form nonlocal investors. This could also be the case with natural disaster risk. If regional natural hazard risk increases more than the national average, local investors' risk aversion could increase. Consequently, local investors will have to sell their local stocks to reduce their exposure to risky stocks, causing the local stock prices to decline. Nonlocal investors might view this decline of prices as an arbitrage opportunity to exploit, therefore the predictability will decline over time.

If mispricing based on local risk aversion is the explanation for the results, the natural hazard risk variable should have no effect on future cash flow or dividend growth rates for the affected firms. Besides, it is to be expected that the results will be corrected by nonlocal investors. Therefore the effect is expected to decline over time and to be stronger among companies with lower visibility and high local ownership. This provides an interesting area for future research.

6 Summary and Conclusion

This research investigates the link between regional natural hazard risk and the predictability of excess returns. Even though there is a significant amount of literature on the effects of natural disasters to the economy, no literature has been written about natural disasters in relation to explaining future stock returns. To test the strength and significance of the effects I use Fama and Macbeth (1973) regressions and quintile portfolio sorts.

The results of the FM cross-sectional regression while controlling for standard control variables show a regression coefficient of 0.001 of the natural logarithm of natural hazard risk. In economical terms the difference between the highest and the lowest natural hazard risk quintile yields an average increase of expected excess return of 0.304%. The corresponding significance level is 5%. The predictability of these results is not explained by geographic or industry momentum. The results are also robust to excluding the financial sector or penny stocks from the sample. Different sample periods show that the effect was most present during the years of the financial crisis from 2008 to 2011 and seem to disappear between 2011 and 2014.

The portfolio sorts do not show any results that are statistically distinguishable from zero in the long – short portfolio. Portfolios with high natural hazard risk are more exposed to the market, but are negatively exposed to the HML and UMD factor. By adding profitability and investment factors to the time-series regression, the *t* – statistic corresponding to Jensen's alpha is significant at 2.78. The difference in average monthly excess return between the high and low risk portfolio is 0.360% for the equal weighted portfolio and 0.577% for the value weighted portfolio.

Although this research reveals that there is a possible effect of natural hazard risk in relation to excess stock returns, I believe future studies might improve on this thesis. Because the effect of natural hazard risk to excess stock return does not exceed a *t* – statistic of 3.0 in any of the tests, the effect I find can not be considered as a factor in explaining the cross-section of stock returns. This research is based on historical data to calculate natural hazard risk. It would be interesting to see whether the predictability of the model improves as meteorological risk estimations improve. In additional tests I find that the effect I find is stronger amongst difficult to arbitrage stocks which could indicate that mispricing might cause of the underlying effect. It would be interesting to see further investigation in whether the effect of natural hazard risk on excess returns could be caused by mispricing based on local risk aversion. Also, further research could be done to the cause of the strong negative exposure to the HML and RMW portfolios.

7 Bibliography

Albala-Bertrand, J. M. (1993). Natural disaster situations and growth: A macroeconomic model for sudden disaster impacts. *World Development*. https://doi.org/10.1016/0305-750X(93)90122-P

Annan, K. (2003). Cumbre Mundial sobre la Sociedad de la Información. *Discurso Inaugural de La Primera Fase de La WSIS*.

Becker, B. (2007). Geographical segmentation of US capital markets. *Journal of Financial Economics*. https://doi.org/10.1016/j.jfineco.2006.07.001

Bender, M. A., Knutson, T. R., Tuleya, R. E., Sirutis, J. J., Vecchi, G. A., Garner, S. T., & Held, I. M. (2010). Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. *Science*. https://doi.org/10.1126/science.1180568

Bhootra, A. (2011). Are momentum profits driven by the cross-sectional dispersion in expected stock returns? *Journal of Financial Markets*. https://doi.org/10.1016/j.finmar.2010.12.002

Bias, P. V. (1992). REGIONAL FINANCIAL SEGMENTATION IN THE UNITED STATES. *Journal of Regional Science*. https://doi.org/10.1111/j.1467-9787.1992.tb00189.x

Birkmann, J. (2006). *Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies*. New York: United Nations University.

Birkmann, J. (2007). Risk and vulnerability indicators at different scales:. Applicability, usefulness and policy implications. *Environmental Hazards*, *7*(1), 20–31. https://doi.org/10.1016/j.envhaz.2007.04.002

Bourdeau-Brien, M., & Kryzanowski, L. (2017). The impact of natural disasters on the stock returns and volatilities of local firms. *Quarterly Review of Economics and Finance*. https://doi.org/10.1016/j.qref.2016.05.003

Caballeros Otero, R., & Zapata Marti, R. (1995). *The Impact of Natural Disasters on Developing Economies: Implications for the International Development and Disaster Community*.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*.

https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Caron, L. P., Jones, C. G., & Doblas-Reyes, F. (2014). Multi-year prediction skill of Atlantic hurricane activity in CMIP5 decadal hindcasts. *Climate Dynamics*. https://doi.org/10.1007/s00382-013-1773-1

Census Bureau. (2010). *FCC Form 477*.

Chui, A. C. W., Titman, S., & Wei, K. C. J. (2003). Intra-industry momentum: The case of REITs. *Journal of Financial Markets*. https://doi.org/10.1016/S1386-4181(03)00002-8

Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *Proceedings of the National Academy of Sciences, 116*(43), 201907826. https://doi.org/10.1073/pnas.1907826116

CRED. (2019). *The International Disaster Database*.

Dacy, D., & Kunreuther, H. (1969). *The Economics of Natural Disasters: Implications for Federal Policy*. New York: The Free Press.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy*. https://doi.org/10.1086/262098

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

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

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*. https://doi.org/10.1086/260061

García, D., & Norli, Ø. (2012). Geographic dispersion and stock returns. *Journal of Financial Economics*. https://doi.org/10.1016/j.jfineco.2012.06.007

Gromb, D., & Vayanos, D. (2010). Limits of Arbitrage. *Annual Review of Financial Economics*. https://doi.org/10.1146/annurev-financial-073009-104107

Harrigan, F. J., & McGregor, P. G. (1987). INTERREGIONAL ARBITRAGE AND THE SUPPLY OF LOANABLE FUNDS: A MODEL OF INTERMEDIATE FINANCIAL

CAPITAL MOBILITY. *Journal of Regional Science*. https://doi.org/10.1111/j.1467-9787.1987.tb01167.x

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ⋯ and the Cross-Section of Expected Returns. *Review of Financial Studies*. https://doi.org/10.1093/rfs/hhv059

Hochrainer, S. (2009). *Assessing the Macroeconomic Impacts of Natural Disasters Are there Any? The World Bank Sustainable Development Network Vice Presidency Global Facility for Disaster Reduction and Recovery Unit*. Retrieved from http://econ.worldbank.org.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*. https://doi.org/10.2307/2328882

Jegadeesh, N., & Titman, S. (2016). Returns To Buying Winners and Selling Losers. *The Journal of Finance*. https://doi.org/10.1111/1540-6261.00555

Korniotis, G. M., & Kumar, A. (2013). State-Level Business Cycles and Local Return Predictability. *Journal of Finance*. https://doi.org/10.1111/jofi.12017

Lamont, O. A., & Thaler, R. H. (2003). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy*. https://doi.org/10.1086/367683

Lewis, J. (1999). *Development in Disaster-Prone Places: Studies of Vulnerability*. London: Intermediate Technology Publications.

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*. https://doi.org/10.2307/1924119

Moore, C. L., & Hill, J. M. (1982). INTERREGIONAL ARBITRAGE AND THE SUPPLY OF LOANABLE FUNDS. *Journal of Regional Science*. https://doi.org/10.1111/j.1467-9787.1982.tb00772.x

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum? *Journal of Finance*. https://doi.org/10.1111/0022-1082.00146

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

Nakamura, E., Steinsson, J., Barro, R., & Ursúa, J. (2013). Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics*, *5*(3), 35–74. https://doi.org/10.1257/mac.5.3.35

Newey, W. K., & West, K. D. (1987). Hypothesis Testing with Efficient Method of Moments Estimation. *International Economic Review*. https://doi.org/10.2307/2526578

Noy, I., & Nualsri, A. (2011). Fiscal storms: Public spending and revenues in the aftermath of natural disasters. *Environment and Development Economics*, *16*(1), 113–128. https://doi.org/10.1017/S1355770X1000046X

Odeh, D. J. (2002). Natural hazards vulnerability assessment for statewide mitigation planning in Rhode Island. *Natural Hazards Review*. https://doi.org/10.1061/(ASCE)1527-6988(2002)3:4(177)

Parsons, C. A. (2016). Geographic Momentum. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2780139

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*. https://doi.org/10.1086/374184

Pirinsky, C., & Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *Journal of Finance*. https://doi.org/10.1111/j.1540-6261.2006.00895.x

Raddatz, C. (2009). *The Wrath of God Macroeconomic Costs of Natural Disasters*. Retrieved from http://econ.worldbank.org.

Schumacher, I., & Strobl, E. (2011). Economic development and losses due to natural disasters: The role of hazard exposure. *Ecological Economics*. https://doi.org/10.1016/j.ecolecon.2011.09.002

Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Shleifer, A., & Vishny, R. W. (1997). A survey of corporate governance. *Journal of Finance*. https://doi.org/10.1111/j.1540-6261.1997.tb04820.x

Simpson, D. M., & Human, J. R. (2008). Large-scale vulnerability assessments for natural hazards. *Natural Hazards*. https://doi.org/10.1007/s11069-007-9202-6

Smajlbegovic, E. (2019). Regional economic activity and stock returns. *Journal of Financial and Quantitative Analysis*, *54*(3), 1051–1082. https://doi.org/10.1017/S0022109018001126

Statista. (2019). *Dossier: Natural Disasters*.

Stéphane, H. (2014). Natural Disasters and Climate Change. In *Natural Disasters and Climate Change: An Economic Perspective* (Vol. 30). https://doi.org/10.1007/978-3-319-08933-1

Strobl, E. (2008). *THE ECONOMIC GROWTH IMPACT OF HURRICANES: EVIDENCE FROM U.S. COASTAL COUNTIES*. Retrieved from http://www.bea.gov/katrina/index2.htm.

Strobl, E. (2011). The economic growth impact of hurricanes: Evidence from U.S. coastal counties. *Review of Economics and Statistics*. https://doi.org/10.1162/REST_a_00082

T. West, C., & G. Lenze, D. (1994). Modeling the Regional Impact of Natural Disaster and Recovery: A General Framework and an Application to Hurricane Andrew. *International Regional Science Review*.

Tavor, T., & Teitler-Regev, S. (2019). The impact of disasters and terrorism on the stock market. *Jàmbá Journal of Disaster Risk Studies*. https://doi.org/10.4102/jamba.v11i1.534

UNEP, & GRID-Geneva. (2013). *PREVIEW Global Risk Data Platfrom*.

UNU-EHS. (2018). *World Risk Report 2018*. https://doi.org/10.3779/j.issn.1009-3419.2009.09.004

Velásquez, C. A., Cardona, O. D., Carreño, M. L., & Barbat, A. H. (2014). Retrospective assessment of risk from natural hazards. *International Journal of Disaster Risk Reduction*. https://doi.org/10.1016/j.ijdrr.2014.05.005

Wang, L., & Kutan, A. M. (2013). The impact of natural disasters on stock markets: Evidence from Japan and the US. *Comparative Economic Studies*, *55*(4), 672–686. https://doi.org/10.1057/ces.2013.16

Worthington, A. C. (2008). The impact of natural events and disasters on the Australian stock market: A GARCH-M analysis of storms, floods, cyclones, earthquakes and bushfires.

Global Business and Economics Review, *10*(1), 1–10. https://doi.org/10.1504/GBER.2008.016824

Xiao, Y. (2011). Local economic impacts of natural disasters. *Journal of Regional Science*, *51*(4), 804–820. https://doi.org/10.1111/j.1467-9787.2011.00717.x

Zhang, X. F. (2007). Accruals, investment, and the accrual anomaly. *Accounting Review*. https://doi.org/10.2308/accr.2007.82.5.1333

Appendix

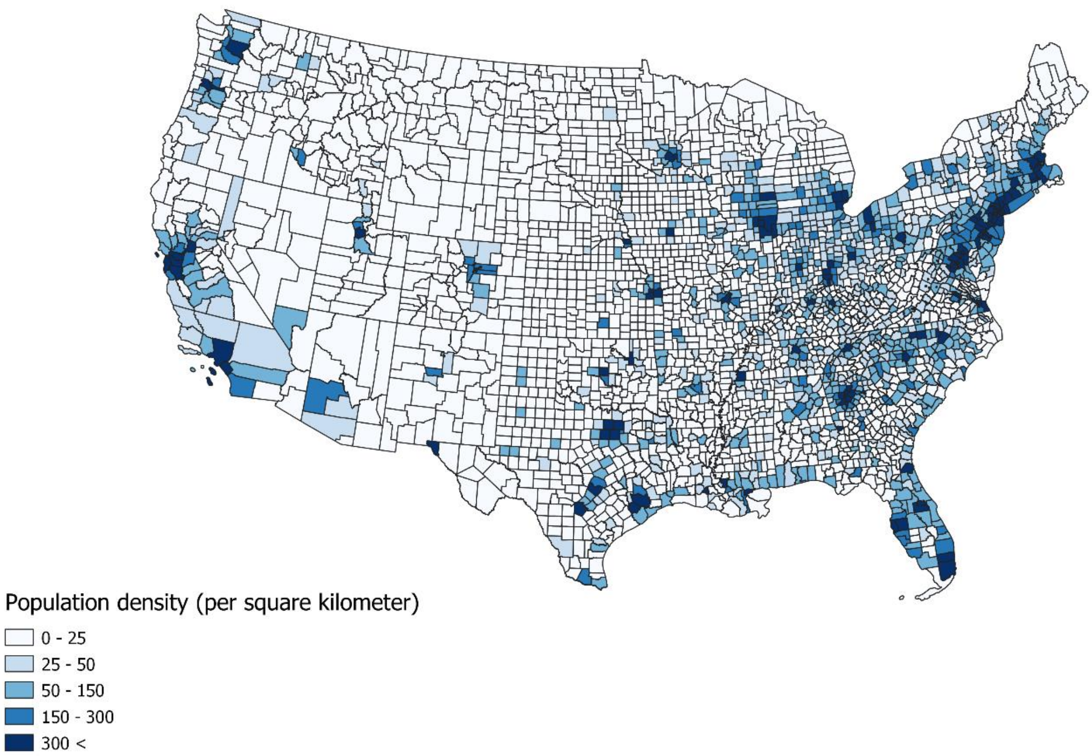
A. Exposure

In this section some additional information and insight on the calculation of exposure to natural disasters is given. Most of this section consists of maps created with QGIS 3.4 software. In this thesis this software has been used to calculate regional exposure to natural disasters.

I chose to calculate exposure based on the relative amount of people affected in a state. Map 1 presents the population density of different counties within the U.S.. It is clear that there is a large distinction between states, but also within different States. To calculate the exposure risk, I use spatial data from GRDP on the exposed area of previous natural disasters per square kilometre using a Graphic Information System (GIS). I combine this data with frequency data from EM-DAT datasets. I combine this spatial data with data from LandScan which provides annual population datapoints per square kilometre.

Of the natural disasters, cyclone surges, cyclone winds and floods are more predictable than earthquakes. This is because these disasters are seasonal. To account for the seasonality of these disasters, I use the average monthly occurrence of the natural disaster based on of year $y - 20$ to $y - 1$ to predict returns of month t in year y.

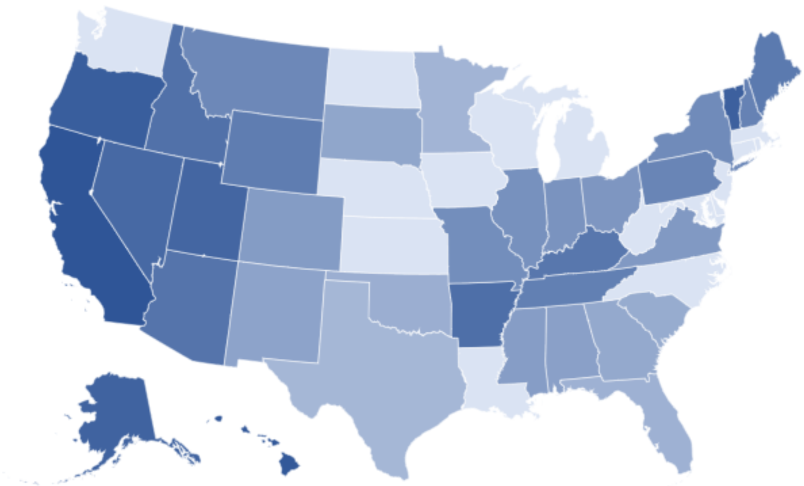
Map 1. Population Density (2017)



To give an indication of the geographical spread of natural disasters I include Map 2 to Map 5. On the U.S. state maps the average annual risk of the specific natural disasters is given. A darker shade of blue indicates a higher risk.

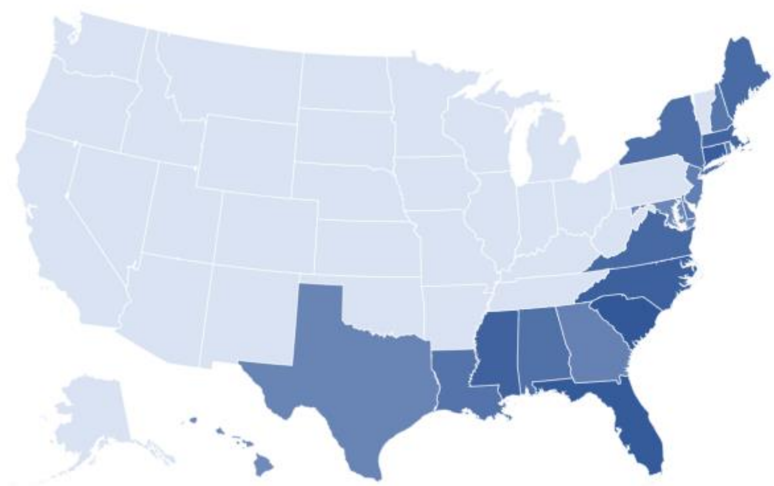
Map 2 reports the earthquake risk. An earthquake is any seismic event generating seismic waves. Most earthquakes are caused by tectonic movement of fault planes. They can also be caused by volcanic activities or landslides etc. but this is less common.

Map 2. Earthquake Risk



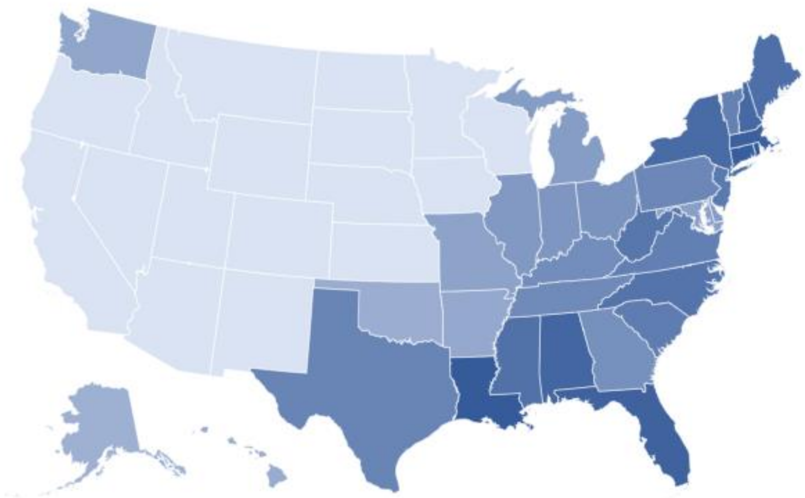
Map 3 reports on cyclone surges. A surge is a sudden rise in sea level in coastal regions caused by tropical storms such as cyclones.

Map 3. Cyclone Surges



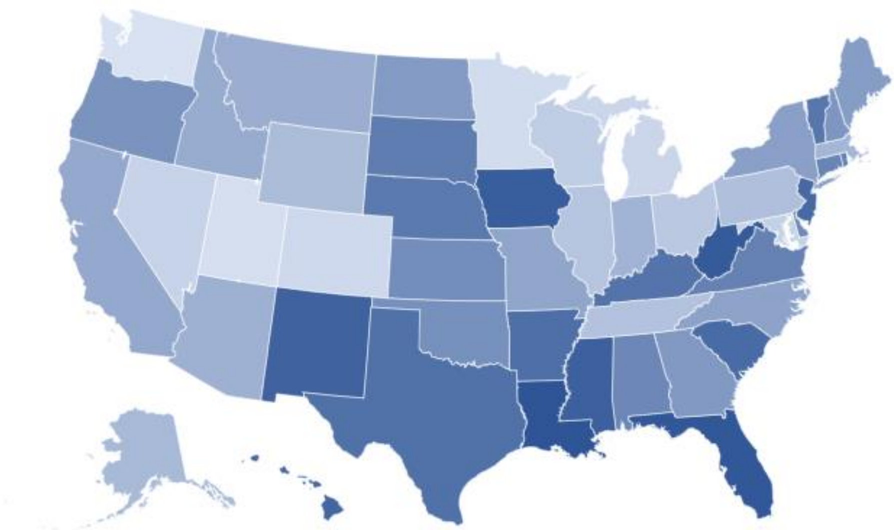
Map 4 presents the U.S. states in which tropical cyclone winds affect the most people (relative to the state's population). The effects of a tropical storm are usually extremely powerful winds combined with heavy rains.

Map 4. Cyclone Winds



Map 5 presents the exposure of floods in different U.S. states. Unlike cyclone surges and winds, the exposure to floods is distributed more evenly across the country. This is because floods are triggered by multiple events. For instance flash floods around the coast line, but also river floods and urban floods caused by heavy rain and/or poor drainage.

Map 5. Flooding



B. Vulnerability

In this section I add some additional information on the calculation of the vulnerability of U.S. states.

To calculate the vulnerability of U.S. states I use the The Social Vulnerability Index (SoVI). The SoVI has been created to help public officials specialized in natural hazards or disease control assess which communities will most likely need help before, during and after a disaster. I calculate the SoVI based on state data. The variables used for constructing the SoVI have changed a bit over the years, but the essence stayed the same. I use the 2018 variables and historical census data for calculating the SoVI between 1990 and 2014.

SoVI exists of 15 variables grouped in four related themes: Socioeconomic status, Household Composition & Disability, Minority Status & Language and Housing Type & Transportation. Figure 1 below shows all the relevant variables and their respective groups.

Figure 1

Variables used in SoVI calculation

<table>
<tr><td rowspan="15">Overall Vulnerability</td><td rowspan="4">Socioeconomic Status</td><td>Below Poverty</td></tr>
<tr><td>Unemployed</td></tr>
<tr><td>Income</td></tr>
<tr><td>No High School Diploma</td></tr>
<tr><td rowspan="4">Household Composition & Disability</td><td>Aged 65 or Older</td></tr>
<tr><td>Aged 17 or Younger</td></tr>
<tr><td>Older than Age 5 with a Disability</td></tr>
<tr><td>Single Parent Household</td></tr>
<tr><td rowspan="2">Minority Status & Language</td><td>Minority</td></tr>
<tr><td>Speaks English ‘Less than Well’</td></tr>
<tr><td rowspan="5">Housing Type & Transportation</td><td>Multi-Unit Structures</td></tr>
<tr><td>Mobile Homes</td></tr>
<tr><td>Crowding</td></tr>
<tr><td>No Vehicle</td></tr>
<tr><td>Group Quarters</td></tr>
</table>

Each of the U.S. states are ranked based on the 15 variables. For each variable a percentile ranking is made. To find the final vulnerability score, the percentile ranks for each variable from each state are summed, and an overall percentile rank is calculated.

C. Correlations

In this section I test for correlations between the natural hazard risk and locality of firms. I find that NHRISK is positively correlated to HHI. Which means that firms with higher natural hazard risk are more locally concentrated. Also, the NHRISK is negatively correlated with STATEDISP, meaning that firms with higher NHRISK are active in less states.

Table 1

Correlation Table

This table presents the correlations between variables NHRISK, HHI and STATEDISP

Variables	NHRISK	HHI	STATEDISP
NHRISK	1.000		
HHI	0.123	1.000	
STATEDISP	-0.119	-0.583	1.000