This thesis draws heavily on the work of Hong and. Al (2016). In their paper, they study if drought risks are priced by markets. To do so, they analyze the actions of food companies in 31 different countries. The drought risk exposure of these companies is approximated by the drought risk exposure of the countries in which they operate. Exposure to drought risk for each country is provided by the Palmer Drought Severity Index (PDSI), which is a widely used metric in climate studies. The authors use this index to predict a drought trend for each country, based on the scientific hypothesis that global warming is responsible for increasing both the frequency and intensity of extreme climatic conditions. Thus, at each given period t, countries are ranked according to their drought trend, which is supposed to forecast their risks of future droughts. Based on these rankings, the authors test whether their prediction allow them to find a statistically significant difference in the returns of stocks of companies located in countries with poor PDSI trends and high PDSI trends. Such a finding would indicate that markets are not efficiently pricing drought risks, since information available at time t allows to predict market movements at date t+1. They take advantage of their predictions of over-valued assets (stocks with low PDSI trend) and under-valued assets (stocks with high PDSI trend) by implementing a long-short strategy that shows abnormal returns over the period 1985-2014. Such finding indeed shows that markets did not price drought risks.

Our wish in this thesis is to extend the findings of Hong et. Al. We consider that the topic of climate physical risks is still very under-represented in the financial literature compared to the multitude of papers that addressed carbon risks and the hype around ESG investing in the past few years in the industry. This is concerning because in a world where a 2° trajectory is not accomplished, physical risks from climate change will undoubtedly take a significantly larger place than carbon risks, which by definition would have not been severe enough. Unfortunately, considering the latest scientific findings and the policies adopted by governments in response, a 2° trajectory seems more and more unlikely. Therefore, our goal in this thesis is to assess to what extent natural disasters, whose intensity and frequency are likely to increase in the coming years, can affect the value of financial assets. This assessment is implemented through an empirical study. We adopt a similar methodology than the one of Hong et Al.

First, we identify a natural disaster whose impact on a specific sector of companies is well-defined and unambiguous. For this regard, we consider storms and electric utility companies. Storms can indeed cause significant financial losses to electric utilities by destroying their installation which typically are outside, in streets and public areas. Second, we need to be able to segregate companies with respect to their exposure to this risk. In the case of a natural disaster, this can only be done through geographical segmentation. Hence, we consider electric utilities that are present on all the states of the US. Unfortunately, an index such as the PDSI for storms does not exist. We thus decide to perform our own measurement of states exposure to storms by using the database of the National Climatic Data Center (NCDC). It provides the historical of all storms that occurred on US soil from 1950 to 2020. The database gives detail on the precise location and date of the storm, as well as its magnitude, the human losses associated to it, and the estimated financial losses. We decide to create our storm risk measure based on the financial losses caused by storms. In the database, this comes in two form : property damages and crops damages. For each storm event, we sum the two and obtain a total estimated financial loss for each storm. Third, the goal is to be able to have a relative measure of storm risks across states. In other words, we would like to assess what states are more exposed to storm risks than others. To do so, we segregate our initial database with respect to the state where the storms took place. At this point, for each US state, we have data describing each storm that happened between 1950 and 2019, with the estimated associated financial loss. Fourth, we aim at verifying if our initial hypothesis holds, i.e. if storms impact the financial performance of electric utility companies. To do so, we rank, each year between 1995 and 2019, the US states depending on how much financial loss they experienced from storm. We form four groups of states depending on their storm exposure, going from the less exposed states to the most exposed. We then segregate the companies depending on their belonging to each group, and we test if, on average, on an annual basis, companies belonging to less exposed groups present better financial performance than companies belonging in more exposed groups. Finally, we test the market efficiency hypothesis. To assess whether markets anticipate storms-related risks, we verify if information regarding storms risks at date t can forecast stock returns of the concerned companies at date t+1. More particularly, we verify market efficiency following Hong et. Al, by assessing if a long-short strategy on supposedly under- and over-exposed stocks to storm risks yields abnormal returns. To do so, we must first create a risk measure which would allow us to forecast future risks. Hong et. Al choose to use an AR(1) model augmented with a deterministic time trend component to estimate future PDSI. We tried to fit a similar model to the annual financial losses from storms but it yielded very poor predictive power. Instead, we decide to test our long-short strategy using two risk measures: (i) we calculate the 99-th percent Value-at-Risk (VaR) of each state up to a given year and we use it to rank the states for the next year ; (ii) we perform a Mann-Kendall trend test for the losses of each state and rank them according to the magnitude of this trend calculated by the Sen’s Slope. The VaR is particularly suited for this kind of exercise, as our data consists of a time series of losses; hence, calculating the VaR at a given level of confidence $\alpha$ simply consists in calculating the $\alpha$-percentile of the time series. This is the so-called historical approach to compute VaR, where the assumption is that future risk can be estimated from past data. We will see if this assumption holds empirically in this case. Using the Mann-Kendall trend test allows us to account for the climate assumption that global warming tends to produce an increasing trend to natural disasters. Using this test is also particularly suited in this case, as a trend is supposed to capture a long-term relationship, that can potentially hold in the future. Countries are thus ranked according to the magnitude of their trend, which is given by the Sen’s Slope. For each method, we then test if the long-short strategy implemented the following year yields abnormal positive returns. Such a finding would indicate a potential mispricing of climate risks.

Pour vraiment isoler l’effet de la zone géographique sur les stock returns, on compare les rendements de stock du même secteur qu’on différencie selon la zone géographique. Mais cela ne montre pas de grandes différences.

On peut alors choisir de comparer les rendements d’actions, sans différencier par secteur, en différenciant par le temps, c’est à dire en comparant le rendement des actions dans différentes zones géographiques avant et après des catastrophes naturelles majeures.

1. Identifier le top 10 des catastrophes naturelles ayant eu lieu aux US entre 1996
2. Comparer la perf moyenne des entreprises se situant dans la zone avant et après chaque catastrophe.

\newpage

\section{2\hspace{0.25cm} Methodology and Data}

\subsection{Overview}

This thesis draws heavily on the work of Hong et. al (2016). In their paper, they study if drought risks are priced by markets. To do so, they analyze the stock returns of food companies in 31 different countries. The drought risk exposure of these companies is approximated by the drought risk exposure of the countries in which they operate. Exposure to drought risk for each country is provided by the Palmer Drought Severity Index (PDSI), which is a widely used metric in climate studies. The authors use this index to predict a drought trend for each country, based on the scientific hypothesis that global warming is responsible for increasing both the frequency and intensity of extreme climatic conditions. Thus, at each given period $t$, countries are ranked according to their drought trend, which is supposed to forecast their risks of future droughts. Based on these rankings, the authors test whether their prediction allow them to find a statistically significant difference in the returns of stocks of companies located in countries with poor PDSI trends and high PDSI trends. Such a finding would indicate that markets are not efficiently pricing drought risks, since information available at time $t$ allows to predict market movements at date $t+1$. They take advantage of their predictions of over-valued assets (stocks with low PDSI trend) and under-valued assets (stocks with high PDSI trend) by implementing a long-short strategy that shows abnormal returns over the period 1985-2014. Such finding indeed shows that markets did not price drought risks. \\

Our wish in this thesis is to extend the findings of Hong et. al. We consider that the topic of climate physical risks is still very under-represented in the financial literature compared to the multitude of papers that addressed carbon risks and the hype around ESG investing in the past few years. This is concerning because in a world where a +2 degrees trajectory is not accomplished, physical risks from climate change will undoubtedly take a significantly larger place than carbon risks, which by definition would have not been severe enough. Unfortunately, considering the latest scientific findings and the policies adopted by governments in response, a +2 degrees trajectory seems more and more unlikely. Therefore, our goal in this thesis is to assess to what extent natural disasters, whose intensity and frequency are likely to increase in the coming years, can affect the value of financial assets. This assessment is implemented through an empirical study. We adopt a similar methodology than the one of Hong et. al. \\

First, we identify a natural disaster whose impact on a specific sector of companies is well-defined and unambiguous. For this regard, we consider storms and electric utility companies. Storms can indeed cause significant financial losses to electric utilities by destroying their installation which typically are outside, in streets and public areas. Second, we need to be able to segregate companies with respect to their exposure to this risk. In the case of a natural disaster, this can only be done through geographical segmentation. Hence, we consider electric utilities that are present on all the states of the US. Unfortunately, an index such as the PDSI for storms does not exist. We thus decide to perform our own measurement of states exposure to storms by using the database of the National Climatic Data Center (NCDC). It provides the historical of all storms that occurred on US soil from 1950 to 2019. The database gives detail on the precise location and date of the storm, as well as its magnitude, the human losses associated to it, and the estimated financial losses. We decide to create our storm risk measure based on the financial losses caused by storms. In the database, this comes in two form : property damages and crops damages. For each storm event, we sum the two and obtain a total estimated financial loss for each storm. Third, the goal is to be able to have a relative measure of storm risks across states. In other words, we would like to assess what states are more exposed to storm risks than others. To do so, we segregate our initial database with respect to the state where the storms took place. At this point, for each US state, we have data describing each storm that happened between 1950 and 2019, with the estimated associated financial loss. Fourth, we aim at verifying if our initial hypothesis holds, i.e. if storms impact the financial performance of electric utility companies. To do so, we rank, each year between 1996 and 2019, the US states depending on how much financial loss they experienced from storm. We choose 1996 as the starting year of our analysis to have consistency in our data, as the data before that year only captured 4 types of storm events, while the data after captures 48 types of storm events. We form four groups of states depending on their storm exposure, going from the less exposed states to the most exposed. We then segregate the companies depending on their belonging to each group, and we test if, on average, on an annual basis, companies belonging to less exposed groups present better financial performance than companies belonging in more exposed groups. Finally, we test the market efficiency hypothesis. \\

To assess whether markets anticipate storms-related risks, we verify if information regarding storms risks at date $t$ can forecast stock returns of the concerned companies at date $t+1$. More particularly, we verify market efficiency following Hong et. al, by assessing if a long-short strategy on supposedly under- and over-exposed stocks to storm risks yields abnormal returns. To do so, we must first create a risk measure which would allow us to forecast future risks. Hong et. al choose to use an $AR(1)$ model augmented with a deterministic linear time trend component to estimate future PDSI. We tried to fit a similar model to the annual financial losses from storms but it yielded very poor predictive power. Instead, we decide to test our long-short strategy using two risk measures: (i) we calculate the 99\% Value-at-Risk (VaR) of each state up to a given year and we use it to rank the states for the next year; (ii) we perform a Mann-Kendall trend test for the losses of each state and rank them according to the magnitude of this trend calculated by the Sen’s Slope. The VaR is particularly suited for this kind of exercise, as our data consists of a time series of losses; hence, calculating the VaR at a given level of confidence $\alpha$ simply consists in calculating the $\alpha$th-percentile of the time series. This is the so-called historical approach to compute VaR, where the assumption is that future risk can be estimated from past data. We will see if this assumption holds empirically in this case. Using the Mann-Kendall trend test allows us to account for the climate assumption that global warming tends to produce an increasing trend to natural disasters. Using this test is also particularly suited in this case, as a trend is supposed to capture a long-term relationship, that can potentially hold in the future. Countries are thus ranked according to the magnitude of their trend, which is given by the Sen’s Slope. For each method, we then test if the long-short strategy implemented the following year yields abnormal positive returns. Such a finding would indicate a potential mispricing of climate risks.

\subsection{Data}

We initially import two types of data: climatic data and financial data. \\

\subsubsection{Climatic data}

The climatic data comes from the National Climatic Data Center (NCDC), a global meteorological data collection centre operated by the National Oceanic and Atmospheric Administration (NOAA). It consists of extreme climatic events that occurred in the United States, between 1950 and 2020, for a total of $N = 925'229 $ observations. Each observation contains more than 30 variables that describe the type of natural disaster, its date and location, the human and the estimated financial losses associated to it, as well as other variables describing the intensity, the source, etc.

We should be careful when using this data as it presents heterogeneity across time. As illustrated on Fig. \ref{Timeline}, a screenshot from the official website of the NCDC \footnote{https://www.ncdc.noaa.gov/}, the type of events collected in the data varies across time.

\begin{figure}[H]

\centering

\includegraphics[width=1\linewidth]{Methodology and Data/EventsTimeline.png}

\caption{Event types timeline. Source: NCDC}

\label{Timeline}

\end{figure}

Until 1996, the events captured are only related to storms, tornadoes and hurricanes. The data becomes much more complete in 1996, covering 48 types of natural disasters, such as storms, droughts, heatwaves, floods, tsunamis, wildfires, and so on. The distribution of the types of natural events covered in the dataset is shown on Fig. \ref{EventDist}.

\begin{figure}[H]

\centering

\includegraphics[width=0.9\linewidth]{Methodology and Data/EventsDistribution.png}

\caption{Event types distribution in total dataset (\%)}

\label{EventDist}

\end{figure}

According to the National Weather Service Instruction (NWSI) \footnote{https://www.ncdc.noaa.gov/stormevents/pd01016005curr.pdf}, as of 1996, the data documents: \\

\textit{a. The occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce;} \\

\textit{b. Rare, unusual, weather phenomena that generate media attention, such as snow flurries in South Florida or the San Diego coastal area; and} \\

\textit{c. Other significant meteorological events, such as record maximum or minimum temperatures or precipitation that occur in connection with another event.} \\

For our analysis, we are interested in assessing the climate-related financial risk for each federal state. The dataset initially contains the losses caused by each natural disaster in terms of property and crops damages. These variables, expressed in US Dollars, must however be considered for what they are: estimates. Moreover, the completeness of this data naturally comes with a trade-off of liability, as multiple sources were used when collecting information. Hence, a strong focus should be put on the detection of outliers. We first focus on detecting outliers associated to financial losses.

Initially, a different row in the dataset accounts for a different date, but the same event can be separated in multiple rows because it had impacts on different dates. Each event being defined by a specific identification, we first group the data by ID event. Then, we look for the costliest natural disasters. One caveat here is that the estimated damages do not take into account inflation. Nevertheless, we observe that the costliest natural disaster in our data is Hurricane Harvey, a category 4 hurricane that costed more than \$USD 38.2 billion in terms of property damages. This is consistent with the report of the NCDC which lists the most expensive weather events that happened in the US between 1980 and 2020. \footnote{https://www.ncdc.noaa.gov/billions/events.pdf} When searching for outliers associated to financial losses further more, we do not find any suspiciously excessive value. However, the column clearly contains some typos as some value are not numbers, or are missing: we remove any observation that fits this criteria. We also look for suspicious values in the Number of Deaths column, and we do not find any excessive nor negative value. At the end, the relevant variables for our analysis are:

\begin{itemize}

\item Beginning date of the disaster

\item Location of the disaster (State)

\item Type of disaster

\item Estimated financial loss from property damages

\item Estimated financial loss from crops damages

\end{itemize}

\subsubsection{Financial data}

We import two financial datasets from the Wharton Research Data Services (WRDS). Both datasets are initially gathered by by Compustat North America.

The first dataset provides fundamental information from all available US companies, between 1961 and 2019. We decide to import the quarterly net income and total assets to calculate the Return on Asset (ROA) of each company. The ROA is a commonly used metric to measure the profitability of a company, and it is particularly suited to compare the performance of companies within the same sector. Companies are identified through a GVKEY, which is a unique identification code for each company. To differentiate between sectors, we import the NAICS codes. The geographical location of each company is also available through the country and state, if available, of their headquarters. Finally, we also import the outstanding number of shares as well as the closing stock price, for each company, on a quarterly basis. This allows us to calculated the Market Capitalization of the companies. This is particularly useful as we would like to make sure that the geographical location of each company corresponds to the location where they operate. Hence we disregard any company with a market capitalization larger than \$1 billion, thus ensuring that we are not dealing with multinational companies. We also disregard any observation for which the absolute value of the quarterly ROA exceeds 500\%. Our first financial data set thus contains fundamental data on 2372 American low and medium sized companies in the sectors of Agriculture (11), Energy (22) and Construction (23), between 1970 and 2019. The numbers in brackets correspond to the NAICS codes of the respective sectors. The number of companies per federal state is given in Table \ref{tab:NbCompanies}.

\begin{tabular}{lr}\caption{Number of companies per US state}

\toprule

State & Number of companies \\

\midrule

CALIFORNIA & 258.0 \\

TEXAS & 225.0 \\

NEW YORK & 172.0 \\

FLORIDA & 155.0 \\

NEW JERSEY & 72.0 \\

COLORADO & 60.0 \\

ARIZONA & 59.0 \\

ILLINOIS & 59.0 \\

MASSACHUSETTS & 57.0 \\

OHIO & 53.0 \\

PENNSYLVANIA & 52.0 \\

NEVADA & 50.0 \\

GEORGIA & 47.0 \\

MARYLAND & 40.0 \\

MINNESOTA & 40.0 \\

VIRGINIA & 40.0 \\

TENNESSEE & 37.0 \\

CONNECTICUT & 36.0 \\

NORTH CAROLINA & 30.0 \\

MICHIGAN & 28.0 \\

MISSOURI & 27.0 \\

INDIANA & 27.0 \\

WASHINGTON & 26.0 \\

OKLAHOMA & 22.0 \\

KENTUCKY & 20.0 \\

OREGON & 18.0 \\

WISCONSIN & 17.0 \\

UTAH & 17.0 \\

HAWAII & 15.0 \\

IOWA & 14.0 \\

LOUISIANA & 12.0 \\

SOUTH CAROLINA & 11.0 \\

DELAWARE & 11.0 \\

KANSAS & 10.0 \\

\bottomrule

\label{tab:NbCompanies}

\end{tabular}

The second dataset contains market information from all available US companies between 1961 and 2019. We import the daily stock prices