Politecnico di Milano
SCHOOL OF INDUSTRIAL AND INFORMATION ENGINEERING
Master of Science – Energy Engineering

Storm is brewing: Extreme weather events and firm performance in Europe

Supervisor
Prof. Vincenzo Butticé

Co-Supervisor
Prof. Roberto Bianchini
Prof. Annalisa Croce

Candidate
Tiziano Milazzo – 912956

Academic Year 2019 – 2020

Acknowledgements

I primi che meritano di essere ringraziati per questo lavoro sono senza dubbio alcuno Vincenzo, Roberto e Annalisa. Dimostrando eccezionale disponibilità e comprensione, mi hanno accompagnato lungo questo percorso permettendomi di inoltrarmi in aspetti e argomenti che pensavo prima al di fuori della mia portata. Con una punta di rammarico, non proseguirò il percorso come dottorato, ma costruire questa tesi con voi è stato forse l'esperienza più soddisfacente all'interno del mio percorso al Politecnico.

Grazie al Politecnico di Milano per avermi fornito una formazione così adattabile nell'affrontare i problemi e una serie di mezzi e strumenti di prima qualità, non scontati nel panorama universitario.

Grazie alla mia famiglia, a mia madre e mio padre che mi hanno sempre sostenuto e permesso di completare il mio percorso di studi, senza mai smettere di incoraggiarmi a coltivare le mie passioni.

Grazie a mia madre che mi ha ascoltato quando avevo bisogno di volare lontano, almeno con la fantasia, ed evadere dal Politecnico verso i miei amati e nostalgici lidi ellenici. Sempre entusiasta di accompagnarmi anche nelle questioni più spinose e burocratiche e capace di fornirmi un granitico riferimento di vita.

Grazie a mio padre per i pranzi, le cene, le merende e le spaghettate di mezzanotte. Per l'incondizionato amore dell'arte che riesce ad emanare e trasmettere con le sue parole e per avermi permesso di tenere viva questa mia passione al di fuori del percorso di studi.

Grazie a zia Mara per essermi stata così vicina ed aver stimolato e condiviso con me la passione per il cinema e la lettura. Per essere stata riferimento indiscusso nella costruzione della mia libreria che, nonostante ancora sfiguri di fronte alla sua, comincia ad essere di tutto rispetto.

Grazie a zio Rinaldo e zia Simona per avermi introdotto alle bellezze della montagna. Mi avete acceso una scintilla che, nonostante sia un focherello in nuce, sto gelosamente

coltivando. Un grazie speciale a IL zio poi per essere stato fondamentale riferimento nel plasmare molti aspetti della mia personalità.
Grazie a nonno Giuseppe per aver condiviso con me memorie di un'Italia ormai passata ma viva di un fascino eterno, vivaci affreschi di vita di campagna, angherie fasciste e piccoli personali atti di resistenza. Grazie per avermi insegnato la vitalità e l'arte insiti in un artigianato di qualità. Purtroppo nei giorni scorsi ci hai lasciato e non potrai goderti la proclamazione, ma la tua eredità rimane nei solchi scavati nel legno, nel profilo di mamma e nel tuo ricordo. A te dedico questo lavoro.

Grazie alla LIPU per avermi accompagnato in anni cruciali della mia formazione, per avermi aperto gli orizzonti sul mondo e per avermi permesso di far parte di questa meravigliosa comunità.

Grazie infine a Giorgio, l'amico di una vita. L'unica persona che desidererei al mio fianco in praticamente qualunque occasione, tranne che a tavola quando è il mio turno a pagare. Non mi sono mai sentito così a mio agio come su quel posto da passegero. Spero che tu possa accompagnarmi alla volta del resto di mondo che ci resta da vedere.

Cesano Maderno, 17/3/2021

Sommario

Il presente lavoro di tesi offre un contributo all'esistente letteratura riguardante gli impatti di eventi meteorologici estremi sull'economia. In particolare, la presente analisi si concentra sulle conseguenze di questi eventi sulla performance economica d'impresa, nello specifico riferendosi alle piccole e medie imprese dell'area europea.

Due differenti banche dati sono state costruite per lo svolgimento dell'analisi. La prima contenente informazioni finaziarie da più di otto milioni di imprese, il 95% delle quali piccole e medie, nel periodo 2009-2018. Il secondo composto dalle informazioni su più di 170 eventi climatici estremi che hanno colpito l'Europa nel periodo tra il 2008 e il 2018.

Le informazioni di queste banche dati sono state poi combinate attraverso l'uso della classificazione regionale NUTS3 e l'implementazione di un modello Difference-in-Difference. Il modello ha misurato gli effetti dell'esposizione ad alluvioni, tempeste e incendi su asset totali, ricavi operativi, costi operativi e rendimento degli asset, controllando per età dell'impresa, settore industriale ed effetto annuo.

I coefficienti DID significativi (almeno al 10%), stimati dall'analisi per tutti gli eventi singolarmente, mostrano un effetto medio complessivamente negativo per asset totali, ricavi operativi e costi operativi, mentre il rendimento degli asset si distingue per un effetto medio moderatamente positivo. Ad ogni modo, questi trend medi derivano da una distribuzione di coefficienti di stima perlopiù bilanciata tra effetti negativi e positivi, che fa luce su una reazione non univoca delle variabili rispetto ai fenomeni in esame.

Inoltre, tre eventi campione (uno per ogni tipo) sono stati osservati singolarmente come casi studio. Tutti hanno riportato coefficienti di DID significativi all'1% e negativi per tutte le variabili in esame.

Abstract

The current thesis work offers a contribution to the existing literature regarding the impacts of extreme weather events on economy. In particular, the present analysis focuses on the consequences of these events on firm level economic performance, addressing SMEs in European area.

To perform the analysis, two datasets were built. The first containing ten years frequency (2009-2018) balance sheet concerning financial information from more than eight million companies, more than 95% of which are small and medium sized enterprises. The second including extreme weather event information of more than 170 events across Europe in the time period from 2008 to 2018.

Information from the datasets was then combined through NUTS3 regions classification and a Difference-in-Difference model was implemented. The model measured the treatment effect of floods, storms and wildfires over total assets, operating revenues, costs of operations and return over assets, controlling for company age, industrial sector and year effect.

Significant (at least 10% level) DID coefficients from the analysis, estimated for all events singularly, returned an overall negative mean behaviour for total assets, operating revenues and costs of operations while reporting a slightly positive overall effect for return over assets. Anyway, these mean trends come from a distribution of estimates mostly balanced between positive and negative coefficients that shed light on a not unique response direction.

Furtherly, case study events (one for each event category) were observed singularly out of the distribution context. All of them present DID coefficients at 1% significance level with severe and negative estimates values for all variables.

Extended Abstract

The scope of the current work is to study how extreme weather events affect firm performance of European SMEs. Thus, a propaedeutic review of the literature was performed to investigate the existing body of knowledge in the field.

A Systematic Literature Review process was exploited, implementing a framework to better explore the literature on economic responses to extreme weather event exposure. A thorough review of a final selection of articles (41) published on prestigious journals in the field depicted a growing interest on the topic and a consistent lack of firm level impact analyses. Furthermore, among methodologies, the Difference-in-Difference model was found to be consistently present in literature, in particular when assessing extreme weather event impacts over granular subjects (such as regions, counties or companies).

Studies in literature address the impact of extreme weather events on economic activities concerning several different economic levels, subject scales and event types. Wage fluctuactions, labor markets, housing, bank network responses and economic performance are some of the topics investigated in the literature when subjects are affected by extreme weather events. However, only few papers concern company level impacts and all of them retrieve financial data from publicly listed firms.

Furthermore, a marked lack of papers addressing the European area contributes to shed light on the compelling development need in literature towards that direction. For these reasons, this thesis found its novelty in further investigating a mostly uncovered aspect of the literature implementing a DID model to estimate the effect of extreme weather events on firm performance in the European area, exploting company level balance sheets and focusing on SMEs.

To investigate these effects, the construction of two databases was carried out, one for financial firm data and the other including extreme weather event information, starting from raw sources of information. The resulting databases present a level of completeness not publicly available at the best of author's knowledge.

The construction process needed for both datasets involved different sources of information and multiple softwares. Raw financial information concerning years from 2009 to 2018 was retrieved from Orbis platform by Bureau van Dijk exploting automated operations performed through Python scripts. The event data was instead collected starting from the EMDAT database by CRED and then transformed through a series of manual operations performed via Excel. Then, because of the substantial lack of coordinate information, a geocoding procedure was achieved via Python scripts implementing the Bing Maps Location API tool to retrieve lacking company coordinates. A similar procedure was applied also to event location and then coordinates from both databases were furtherly processed via QGIS software to match locations with the NUTS3 region they are contained in. NUTS3 classification was chosen as a recognized and reliable compromise for geographical matching between firms and extreme weather events. Further operations and final tunings of both datasets were performed via STATA software, as well as the data analysis.

Economic impacts were measured observing immediate firm performance indicators. The choice fell on a group of variables defined as follows:

- *Total assets (TA):* logarithmic value of total assets. Specified by company and by year (2010-2018), dimensionless.
- *Operating revenues (OR):* logarithmic value of operating revenues. Specified by company and by year (2010-2018), dimensionless.
- *Cost of operations (CO):* logarithmic value of the difference between operating revenues and EBITDA. Specified by company and by year (2010-2018), dimensionless.
- *Return over assets (ROA):* winsorized value of EBITDA over total assets. Specified by company and by year (2010-2018), dimensionless.

Then, exploiting NUTS3 information, each event from the database was extracted and singularly matched with financial data. A series of regressions, one for each event analysed, was performed, implementing the DID model through the setting of proper dummy variables. The formulation of the regression model is the following:

$$y = \beta_0 + \beta_1 time + \beta_2 A + \beta_3 did + \beta_4 age + \sum_i \beta_{5i} year_i + \sum_j \beta_{6j} nace_j + \varepsilon$$

Where:

- y : the output variable under study *(TA, OR, CO or ROA).*
- *time* : the dummy variable related to time *(T).*
- *A* : the dummy variable representing treatment.
- *did* : the combined variable that identifies the DID effect *(T·A).*
- *age* : the control variable related to the age of companies.
- $year_i$: the *i* control variable out of ten related to the year of observation.
- $nace_j$: the *j* control variable out of twenty-three related to firm's industrial sector.

Among estimation results, only significant (at least 10% level) DID coefficients were considered. Significant DID estimates do not assume a unique sign characterization, but they are essentially balanced between negative and positive signs. However, the overall mean of coefficients reflecting the DID effect on TA, OR and CO is negative for almost all event categories, while on the contrary the overall mean effect for ROA is slightly positive for each event type. These findings shed light on the complex nature of the phenomena under study suggesting further developments for future applications, such as a cross sectoral control or an improved location specification.

Moreover, a case study estimation for each event category was separately observed. The choice was guided by a high media resonance, often partnered with tragic death tolls and severe damages. For these specific cases, estimations showed a significant and severely negative effect for TA, OR and CO while ROA presented slightly lower trends.

The findings of the present thesis contribute to the literature by investigating a still mostly uncovered aspect in the field. Therefore, due to the novel and explorative approach of the current work, several opportunities for improvements and further developments are available.

For instance, starting from the current thesis, future applications could consider improvements such as an extension of the time period considered (2019-2018), adoption of other financial variable indicators, different, and possibly more accurate, criterion for geographical matching.

Keywords

Climate change, extreme weather events, firm performance, Difference-in-Difference, Europe, SMEs.

Parole chiave

Cambiamento climatico, eventi climatici estremi, performance d'impresa, Difference-in-Difference, Europa, piccole e medie imprese.

Table of Contents

Introduction

Human history has a sorrowful familiarity with extreme weather events. Nature's most dreadful threats have always been a constant in nightmares of mankind. So baffling and harmful phenomena that for long times they have been considered consequences of the anger of gods. Memories of natural hazards from the past have been preserved through myths and legends, thus Noah's massive flood became the consequence of God's punishment for mankind's impiety while the several storms that haunted Odysseus's return were the result of Poseidon's rage.

Nowadays, legends and myths belong to the past, but extreme weather events are still raging and a new concern grips the world, the anger of the gods have given way to human driven climate forcing. Climate change is the challenge of the century and anthropogenic forcing is now out of doubt, but the extremely complex pattern which composes the climate system and the lack of a proper number of observations to define a definitive trend have made difficult deciphering the relationship between climate change and the spread varieties of meteorological events. Therefore, while for heat waves and extreme precipitations trends are already mostly delineated, for other categories such as storms, floods, wildfires and droughts the literature is still debating, even if last century trends do not seem to prospect for the best.

Nevertheless, the topic is growing in interest within the literature and significant developments have been achieved during last decades, in parallel with the assessment of impacts on human society. A systematic review of the literature regarding this latter aspect revealed a substantial lack of studies addressing the impacts of extreme weather events on firm level economic performance, in particular in the European area.

Within this context, the current thesis aims to further investigate this aspect by proposing an explorative approach to assess concequences of firm exposure to extreme weather events, focusing on European small and medium sized enterprises (SMEs).

Balance sheet data for more than eight million European firms covering the period from 2009 to 2018 was exploited in the construction of a financial panel dataset to be combined with another dataset concerning extreme weather event information. The integration of these datasets was then used for the quantitative assessment of impacts of extreme weather events on firm performance.

The thesis is structured as follows:

- Chapter 1 describes the Systematic Literature Review (SLR) procedure exploited in defining the research question for the thesis.
- Chapter 2 concerns the description of the methodology adopted in the construction of datasets and in the model definition for the empirical analysis.
- Chapter 3 reports estimated results of the impacts of extreme weather events on firm performance indicators.

Chapter 1 Literature review

Literature review is a fundamental step in research studies. It has the goal to explore the current knowledge about subjects of interest in the scientific community. To do so researchers need to adopt a methodology to find their way through articles and papers.

This chapter explains the methodology adopted and it also presents a review summary with the aim of picturing a scientific background as complete and accurate as possible. Finally, a research question is expressed and elaborated.

The methodology adopted refers to the principles of systematic literature review (SLR). It differentiates from traditional "narrative" review methods in terms of thoroughness, trasperency and replicability. In management studies a "narrative" review is usually adopted, but it has been argued that a SLR approach can help avoiding typical review problems related to author's subjectivity bias improving rigor and reliability (Tranfield, Denyer & Smart, 2003).

For these reasons the SLR is chosen as the reference method for the review of this thesis. All steps and decisions made by the author are accurately described in order to make the process as reproducible as possible. This approach also helps in providing evidence of novelty while systematically defining the research topic background.

The SLR pool of articles was last updated on 30/11/2020. A further check was performed on 20/03/2021, 16 new articles have been added to Scopus but without a significant change of the literature context.

1.1 Research proposal

The scope of this thesis is to probe impacts of extreme weather events on economy and financial institutions at a local scale.

While concerning extreme weather events, special attention was reserverd to climate change influences. Global warming induced temperature increase is supported by a widespred literature background. In contrast, the influence of climate change on extreme weather events, even if recognized, does not draw a unique trend as for temperature. In fact, extreme weather events are influenced in different ways (e.g. severity, frequency) depending on several aspects (e.g. type of event, location). Anyway, this topic will be further examined in the following.

This thesis contributes to extreme weather event literature by assessing the impact of these events on firm performance.

1.2 Review protocol

Establishing a review protocol is crucial to perform properly a SLR. The protocol was shaped following the need of exploring the impacts of extreme weather events on firm performance in the existing literature.

1.2.1 Database of sources

The current literature was investigated via Scopus by Elsevier. Scopus[1] is an abstract and citation database founded in 2004. It is a powerful and widespread platform that allows the user to build very detailed queries, organize results in lists and retrieve a large amount of metadata on sources.

The access to the platform is granted by the Politecnico di Milano to all students.

1.2.2 Protocol queries

To obtain the final pool of studies, a query was structured and progressively restricted step by step. Starting from a very general query on the topic, filters concerning area of interest, relevance of sources and research customized criteria were applied.

The whole process is graphically depicted in the flow diagram in Figure 1.1.

The first general query ("TITLE-ABS-KEY" in Figure 1.1) regarded the keywords chosen to define the topic. The Scopus complete input is reported below:

[1] https://www.scopus.com/

TITLE-ABS-KEY (("extreme weather event*" OR flood* OR drought* OR storm* OR landslide* OR wildfire* OR hurricane* OR "extreme wind*" OR "extreme temperature" OR "heat wave") AND (performance OR impact OR compan* OR firm* OR financ* OR econom*)) …

This first result was then filtered by area of interest to better focus on the research proposal. ("Subject Filter" in Figure1.1):

… AND (LIMIT-TO (SUBJAREA,"MULT") OR LIMIT-TO (SUBJAREA,"ECON") OR LIMIT-TO (SUBJAREA,"BUSI")) …

Finally, the last category filter concerned relevance of sources. In particular, only papers published in prestigious journals were taken into account ("Source Filter" in Figure 1.1). The list of sources was built on the base of SCImago Journal & Country Rank[2], a publicly available portal that includes journals scientific indicators developed from Scopus database information. The first 35 voices of source in each of "Finance", "Business, Management and Accounting" and "Economics and Econometrics" subject areas were included in the list of "relevant" sources. To this list also *PNAS (Proceedings of the National Academy of Sciences)*, *Nature* and *Science* were added as the three most influencing general-science journals according to the Impact Factor (IF) ranking.

Finally, taking into account the overlapping of some journals in the three different rankings, the final pool of sources contains 83 journals (see Appendix for full list).

After this last filter, the final group of articles was composed by 850 voices (updated 30/11/2020).

The second part of the review protocol involved binary questions. All articles in the final filtered group were subjected to these questions via progressive steps.

[2] https://www.scimagojr.com/

The first step concerned a preliminary screening procedure on the base of title-abstract review: titles and, if necessary, abstracts were considered in giving feedbacks to the first two questions.

BQ1) Does the study cover topics in the field of economics, econometrics, management, finance or policymaking?

BQ2) Does the study assess extreme weather event impacts or models?

If at least one of the questions returned a positive answer the paper was included in a secondary pool to be subjected to the third question, otherwise it was discarded. The third question was answered applying a partial review step: introduction, methodology and conclusions are considered. The binary question was the following:

BQ3) Does the study present an original quantitative model about EWE impacts on human activity?

For a positive answer to this question the paper was included in the final pool of articles. The actual full review step was then performed on this final group. All records from the secondary pool not included in this final group were considered as background literature information.

The final pool of selected documents presents 41 articles (see Appendix for full list).

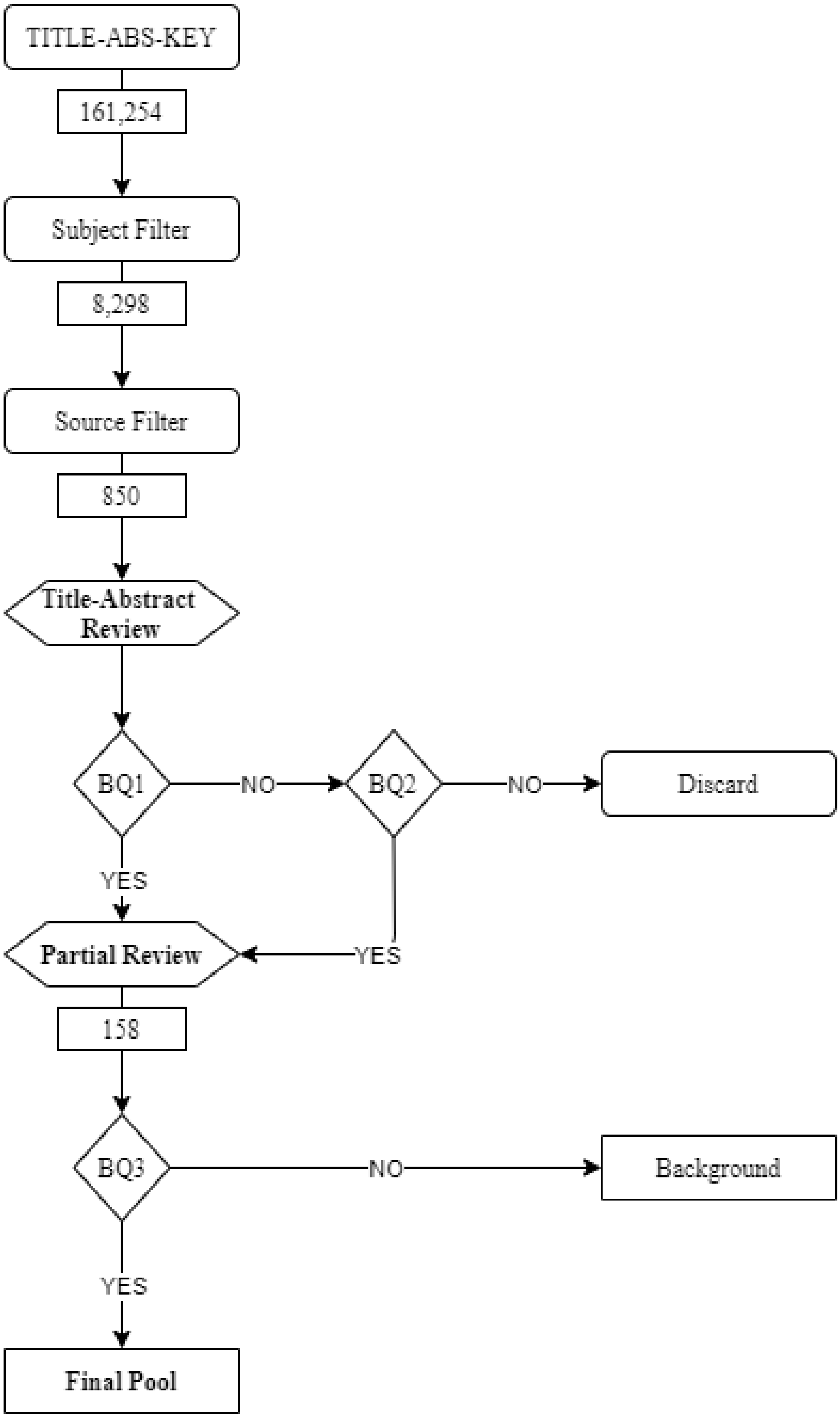

Figure 1.1 – Flowchart diagram of the SLR screening process

1.3 Review description

This section is going to provide an overall outlook of the current literature on EWE influence on human activity. The first part of this section is dedicated to descriptive analysis of the pool through the use of metadata.

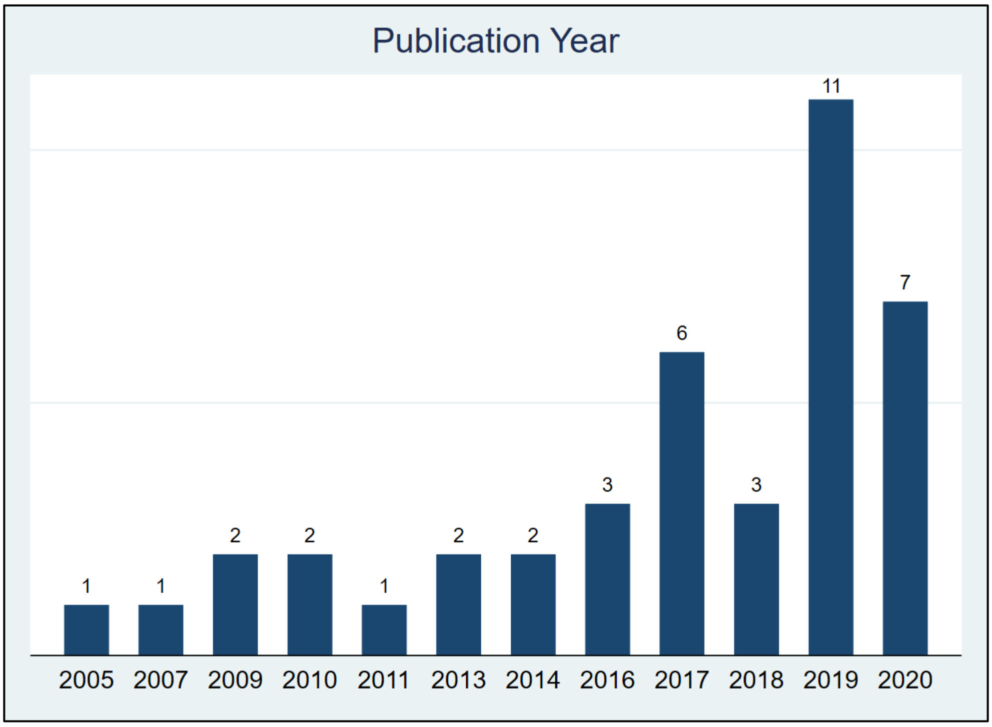

Figure 1.2 – Histograms of number of publications over years from the final SLR pool.

As the graph points out, the portion of scientific literature assessing EWE influence on human activity seems to have raised in interest in the last years. This is probably due to the actual concerning on climate change influence over frequency and intensity of these events: even the oldest article of the set issues the changing climate involved in the drought that affected Europe in 2003, dwelling then on the impact of this event on carbon emission and agricultural production (Ciais, Reichstein et al., 2005). Anyway, the relevant papers for the scope of this thesis all belong to the most recent years of the distribution, underlining how EWE influence on firm level performance is a topic of very actual interest.

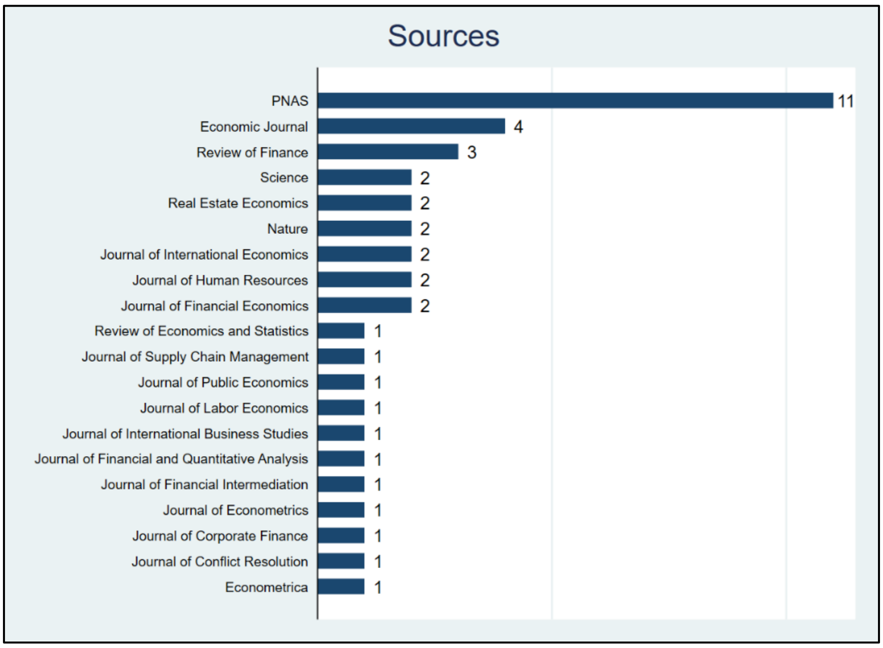

Figure 1.3 – Bar chart of number of publications over sources from the final SLR pool.

It is also interesting to observe the diversified distribution in sources, as evidence of the cross-sectional nature of how these extreme events affect society. Among sources, *Proceedings of the National Academy of Sciences of the United States of America* (PNAS), *Economic Journal* and *Review of Finance* stand out in frequency.

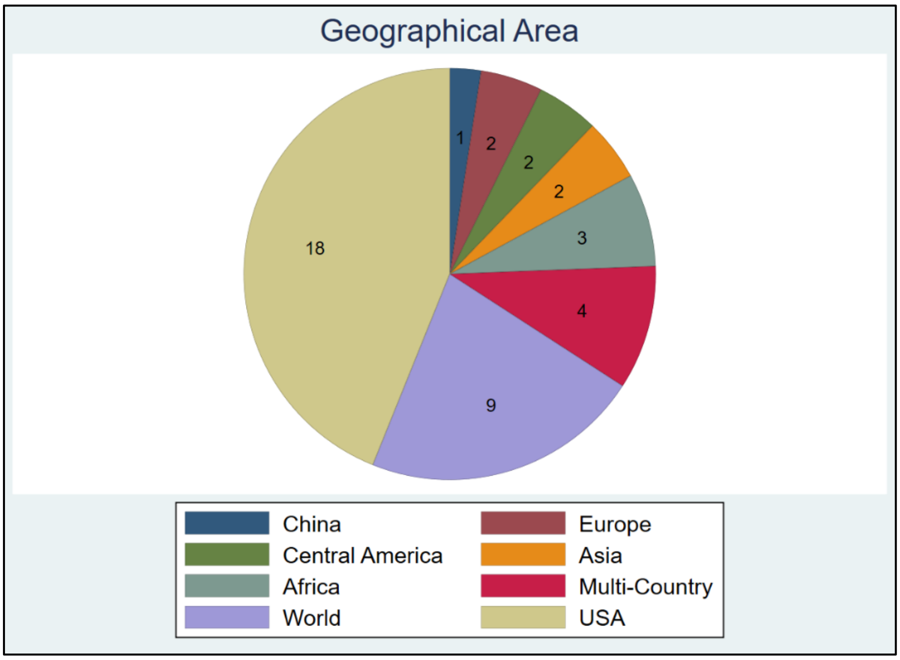

Figure 1.4 – Pie chart of papers from final SLR pool, based on geographical area addressed in the study.

The predominance of PNAS as source is also reflected on both the geographical distribution of countries and the event catagory addressed by the studies: US hurricane impacts on economy is the most common referred topic. It is also relevant to notice how most of the studies not referring to US issues developing countries, consistent with their higher vulnerability to EWE consequences. On the other hand, very few papers focus on Europe, concerning the need of further development.

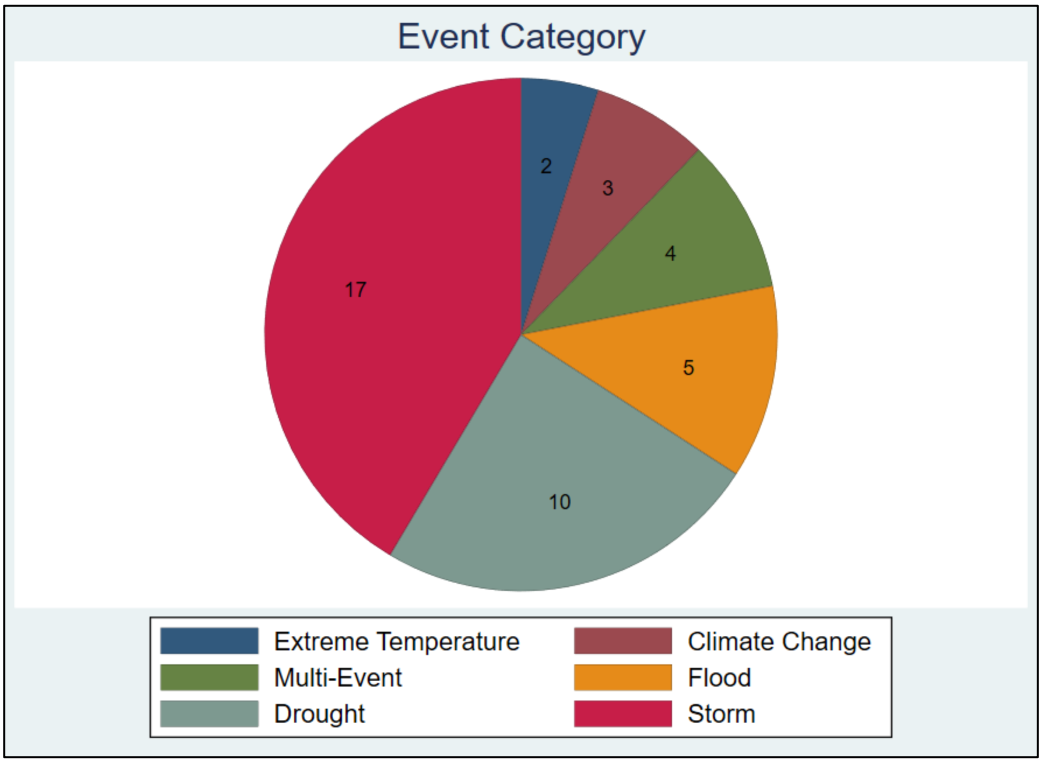

Figure 1.5 – Pie chart of papers from final SLR pool, based on event category addressed in the study.

In the previous graph, frequency of event categories is reported. It may be useful to specify that "Storm" refers to all storm related weather events, such as hurricane and tropical cyclones, while "Climate Change" category is used to identify all papers that assess extreme weather events under study as directly related to the changing climate. As already discussed, storms are the most issued event, followed by drought and flood. No paper was retrieved specifically on wildfire and landslide, even if the former was the main topic addressed in the background literature. In fact, the most of the articles not included in the pool, but saved as background literature, concerns about dreadful wildfire seasons of the last few years.

Finally, it may be interesting to have an overview on the number of citations inside the selected pool. Detailed metadata are provided by Scopus based on "Cited by" information.

Citations Overview

Mean	97
Standard deviation	342
Median	16
Maximum	2193
Minimum	0

Table 1.1 – Summary statistics of Scopus citation metadata from the final SLR pool.

It is immediately visible the presence of an outlier that undermines the correct representation of the overview. The aforementioned article is *"Europe-wide reduction in primary productivity caused by the heat and drought in 2003"* (Ciais, Reichstein et al., 2005) published by *Nature*. It is the oldest article of the pool and it addresses the topic in a very multidisciplinary and cross-sectional way, in addition it is published by a very relevant journal in general science. All these features made it the best candidate to be a probable outlier. Thus, excluding it from the citation analysis, a more representative outlook stands out.

Citations Overview Adjusted

Mean Adj.	45
Standard deviation Adj.	67
Median Adj.	15
Maximum Adj.	269
Minimum Adj.	0

Table 1.2 – Summary statistics from final SLR pool after outlier adjustment.

Consistent with the increasing interest on the topic, the next graph shows the citation trend of articles in the final pool.

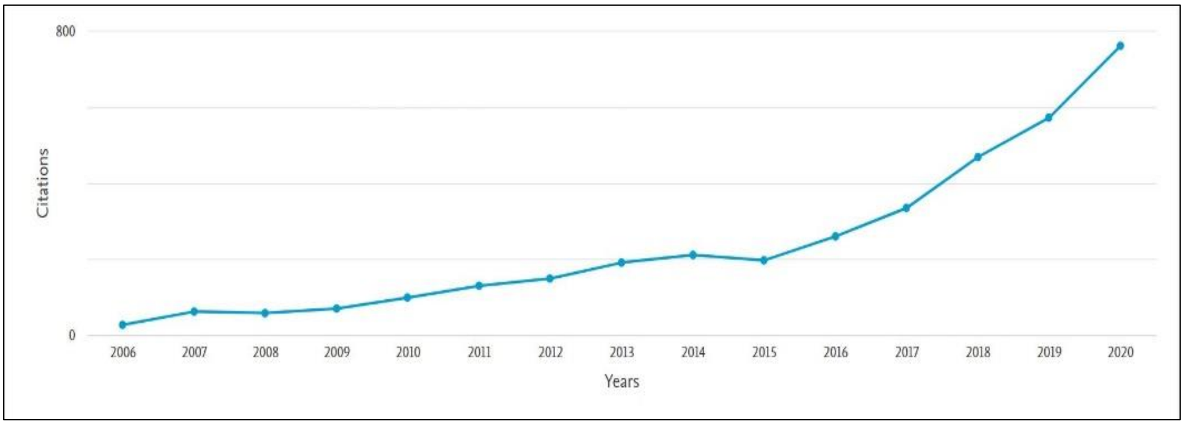

Figure 1.6 – Overall citation trend over years from final SLR pool (image provided by Scopus).

Following subsections will provide a summary of the in detail reviewed articles. There are many criteria that could have been used to describe the present literature on EWE influence on human activity. Extreme weather events are a diversified group of meteorological anomalies and for this reason they impact very differently economy and human life. A leading criterion focused on impacted subject was preferred on a classification based on the nature of shocks. A macro categorization consists in the division between "Non-Economic" articles and "Economic" papers; within these two macro-groups, further classification, based on topics issued, is exploited to organize the articles. Finally, three separated sections are provided: one is dedicated to all, "Economic" and "Non-Economic", articles that were not categorized but that deserve mention, another focuses on the role of *Difference-in-Difference* method among reviewed articles, while the last one addresses climate change influence on extreme weather events.

1.3.1 Non-Economic Literature

As specified before, during the screening procedure a double positive aswers was not necessary to pass the double question step, namely the first question concerning the economic field and the second addressing EWE impacts. For this reason, a relevant, even if minority, portion of the literature assesses non-economic subjects. Within this subgroup of papers three major topics stand out.

1.3.1.1 Healthcare and Social vulnerability

Casualties are one of most relevant consequences of extreme weather events. The death toll during a disaster event is influenced by several factors that are not only related to the severity of the meteorological conditions. In a recent study Tennant and Gilmore issued how government effectiveness, measured in relation with infant mortality rates as an indicator of public service delivery, influences casualties during tropical cyclones. They collected data for over 1000 events from 1979 to 2016 finding that stronger istitutions, independently from income and socioeconomic conditions, lead to a significant reduction in mortality from cyclones (Tennant & Gilmore, 2020).

Social vulnerabilities play a key role in disaster damages and very often the most vulnerables individuals, such as children, are those who pay the higher cost, even in a long run perspective. There is evidence that early childhood extreme temperature and drought exposure have long run repercussions on child stunting (Cooper, Brown et al., 2019), educational attainment (Randell & Gray, 2019) and disabilities rates (Dinkelman, 2017) in several developing countries.

1.3.1.2 Conflict and Social unrest

There is a relevant portion of literature that try to assess the relation of EWE shocks with conflict, violence and social unrest. All the studies agree on the fact that this relationship is not unequivocally true, but that it has to be holistically interpreted to shade light on how these events can influence social unrest.

One of the most recent and discussed episodes on this topic is the Syrian Civil War. In a recent study Ash and Obradovich observed nighttime lights, as a proxy for population density, finding how the worst one in 500 years drought of the Syrian history influenced population displacement. They furtherly interpreted internal migration from most affected provinces as a relevant concurrent cause in raising risk of additional protests in Sunni Arab regions at the time of the 2011 uprising (Ash & Obradovich, 2020).

As for the Syrian Civil War, also other studies agree on the fact that weather shocks can be a relevant, but not sufficient, factor in justifying social unrest: Von Uexkull, Croicu et al. observed how this relation becomes significant only in agriculturally dependent and politically excluded groups in very poor countries (Von Uexkull, Croicu et al., 2016); of the same opinion, Couttenier and Soubeyran highlighted how drought is significantly related to conflict in sub-saharan Africa only for ethnically fractionated countries and with low levels of democracy (Couttenier & Soubeyran, 2014).

These particular social vulnerabilities can easily be found in past societies when social unrest and weather shocks seem to have been more dependent each other. Jia focused on peasant revolts in historical China (between 1470 and 1900) finding how exceptional droughts increased the probability of a revolt, she also showed how the introduction of sweet potatoes contributed to the prevention of drought impact on food supply significantly reducing revolts probability in drought years (Jia, 2014). Another interesting viewpoint on this piece of literature is presented by Chaney: he documented how Nile shocks through centuries influenced the relationship between the sovereign and head religious authority in Islamic Egypt. In fact, he demonstrated how, during these shocks, the religious influence over the sovereign increased because of the very high probability of success of a social revolt coordinated by the head judge (the supreme religious authority) (Chaney, 2013).

1.3.2 Economic Literature

The core of this literature is, for obvious reasons, referred to economic articles, with 26 papers out of 41. They cover a wide range of topics related to EWE impacts on economy: from macro economy to finance, policymaking and firm performance. A summary of this piece of literature is described in the following.

1.3.2.1 Housing and Labor market

Extreme weather events can have a counterintuitive influence on labor markets. On this argument, Belasen and Polachek focused on labor market variations in Florida counties after a hurricane landfall. They found an earnings per worker increase in stricken counties and almost the same decrease for neighboring non-stricken counties. This increase in earnings is also accompanied by a significant employment rate decrease in affected counties. These findings shade light on the complexity of economy in the aftermath of a disaster: in fact, earnings increase because of reconstruction economic efforts but employment data confirm how people often flee away from the disaster zone affecting the labor supply demand curve (Belasen & Polachek, 2009).

Long run effects of extreme temperature and disaster exposure on wages and income are also issued by this portion of literature. Isen, Rossin-Slater et al. exploiting a sample of 30 years old US individuals from 1998 to 2007, discovered a negative correlation between economic outcomes and prenatal extreme temperature days exposure (exceeding 32°C) (Isen, Rossin-Slater et al., 2017). Similarly, Karbownik and Wray focused on prenatal and early life affected US individuals exposed to hurricanes between 1886 and 1897; they observed a 5.7% and a 4.7% lower income during adulthood, in utero and before the age of nine months respectively, with respect to unaffected individuals (Karbownik & Wray, 2019).

Other articles in this section address housing and property markets. Turnbull, Zahirovic-Herbert et al. studied price and liquidity (the inverse of selling time) capitalization in presence of flood risk in East Baton Rouge Parish, Louisiana. They raise concern on how focusing solely on price effects, without considering the liquidity capitalization, could underestimate the real impact on house markets. Furtherly, they show how, in high flood risk zones, the amenity effect (proximity to the waterfront) and insurance benefit underprice do not compensate the negative impact of flood risk on price and selling time (Turnbull, Zahirovic-Herbert et al., 2013). Opposite results are collected by Atreya and Czajkowski when addressing the coastal county of Galveston, Texas: the authors outline how the amenity effect almost completely shadows the flood risk impact leading to a price premium of up to 146% (above average selling price). This finding highlights how, in the presence of a strong amenity effect (Gulf Coast Ocean against Mississippi river in East Baton Rouge) and in the absence of a salient flood disaster, it is very difficult for the actual, objective flood risk to affect people's desire to live next to the Ocean (Atreya & Czajkowski, 2019).

It is worth a special mention to mangroves and coastal wetlands as a very specific topic addressed by some papers in the literature. Sun and Carson, in their articles, examined all tropical storms and hurricanes that affected US coasts between 1996 and 2016. They estimated the Annual Expected Property Damage to assess the impact of shelter reduction against storms due to the consistent coastal wetland ecosystem loss during that period. Then they focused on how much property damage from hurricane Irma (2017) could have been prevented, in the absence of this ecosystem loss: their results returned a potential avoided loss by about $430 million (Sun & Carson, 2020). Finally, two other papers issued the shelter potential of mangrove coastal belts on economic activity in the presence of an extreme event, raising concern on the worring deplection of this habitat in the last decades. Both of them exploited a similar method measuring nighttime lights as an indicator for economic activity: they concluded that mangroves presence is a key factor in reducing long-term permanent economic losses (Hochard, Hamilton and Barbier, 2019) and that a 1km or more of mangrove width is capable to entirely mitigate hurricane effects (del Valle, Eriksson et al., 2020).

1.3.2.2 Banks, Comparative advantage and Policymaking

A very relevant portion of literature addresses finance-related aspects such as lendings/borrowings, risk taking, bonds and stakeholder policymaking.

Schüwer, Lambert and Noth explored how banks react to an exogenous shock, namely Hurricane Katrina, finding that highly capitalized independent banks in the impacted area increased their risk-based capital and strengthened themselves against future liquidity shocks by prioritizing low risk-weighted assets. Differently from banks that are part of a holding company, they invested more in government securities and reduced loan exposures to non-financial firms, while also emitting new lending to them. Affected counties with such bank structure, highly capitalized local independent banks, reported over time higher growth rates in income and employment with respect to other affected counties (Schüwer, Lambert & Noth, 2019). On the other side of the Atlantic, Koetter, Noth and Rehbein performed a similar investigation on the role of local banks during exogenous shocks, in this case Elbe flood in 2013. They measured an increase in lending from exposed local banks, resident in unaffected counties but with interests in stricken firms, with respect to unexposed local banks. They also found that exposed local banks with access to diversified intra-group networks emit *recovery lending* without excessive risk-taking. For this reason, flooded firms without access to such banks experiment decreasing credit after the event (Koetter, Noth & Rehbein, 2020).

Exogenous shocks, especially very extreme events such as Hurricane Katrina, have severe consequences not only on directly affected subjects. Massa and Zhang observed the consequences of the liquidation of bonds holdings driven by property and reinsurance companies after Katrina. They found that the drop in bond prices induces firms to swap from bond financing to bank borrowing, observing also how this change in debt policy tends not to revert even in the long term (Massa & Zhang, 2020). The same spillover effect was previously observed in a different paper by Manconi, Massa and Zhang when studying the relationship between bondholder concentration and credit risk. They essentially outlined a positive relationship between bondholder concentration and credit risk, proxied by corporate bond yield spreads. Bondholder concentration reduces the cost of default (credit risk) inducing firms to undertake riskier policies (Manconi, Massa & Zhang, 2016).

Another evidence of how EWE driven shocks, namely hurricanes, influence policy and investment choices at company level is provided by Pelli and Tschopp. In their article they observed how the destructive power of hurricanes opens the “reconstruction opportunity” of investing in comparative advantage. In particular, such events suddenly reduce the adjusting costs of the production structure inducing industrial companies to invest in state-of-the-art production processes (Pelli & Tschopp, 2017).

In the presence of destructive and emotional impacting calamities, such extreme weather events, perceived risk logics may prevail objective probability of risk. In this regard, Dessaint and Matray conducted a study on how managers react to salient risks, namely liquidity risk. They observed how the managers’ reaction is often disproportionate and not justifiable by the actual extent of risk, evidencing how perceived risk can lead to suboptimal decisions. They bring evidence of a large distortion between perceived and actual cash risk by observing corporate cash holding as a proxy of their risk perception: cash holding increase by companies in the neighborhood of the event is not explainable by the actual probability of an extreme event, suggesting an overreaction to a sudden salient risk. Furthermore, such behaviour is observed only the first and the second time the event occurs, but not successive occasions, and only temporarily, reverting to pre-event levels over the year, being consistent with the theory of the saliency based decision making (Dessaint & Matray, 2017). Futher evidence of exogenous shock influence on managers’ policymaking is provided by Aretz, Banerjee and Pryshchepa in industrial firms risk-shift reaction to hurricane shocks. They highlight how pre-event moderately distressed companies (that they define “ailing”) differently behave with respect to pre-event highly distressed firms. In fact, ailing firms, after the hurricane landfall, reorientate their asset mixes toward abnormally higher-risk segments increasing their probability of failure over the next 10 years, thus hurting creditors (risk-shift). While, on the contrary, pre-event distressed companies adjust their asset mixes toward lower-risk segments. Finally, employing covenant violation data, they observed that distressed firms are more likely to be under creditor control because of covenant violations, while ailing firms often result covenant compliers. This peculiar binding condition prevents distressed firms from risk-shifting by not acting against the interest of creditors, such as not controlled ailing firms do (Aretz, Banerjee and Pryshchepa, 2019).

1.3.2.3 Firm performance

Finally, a relevant piece of literature explores the impact of extreme weather events on firm performance.

Huynh, Nguyen and Truong tested the financial cost of drought on the firm cost of capital by observing expected rate of returns between 1968 and 2015. They documented a higher cost of capital for firms located in states affected by severe drought conditions. They also outlined how this adverse condition is successfully dampened by prevention financial instruments, such as corporate diversification, geographically distributed business operations and cash holdings (Huynh, Nguyen & Truong, 2020). Moreover, Huang, Kerstein and Wang studied how firm performance, proxied by return-on-assets and cash flow, are affected by climate risk through the Global Climate Risk Index, a weighted country level index that assesses for risk from major extreme weather events. As expected, they observed lower and more volatile earnings and cash flows for most exposed countries together with the firms' tendency to have less short-term debt in favour of long-term debt and to distribute lower dividends. In addition, they outline how the climate risk influence on firms varies across industrial sectors depending on several aspects: for examples companies that strongly rely on wide supply chain and infrastructures or industries with non-distributed long-living intensive capital assets are more vulnerable to weather extremes (Huang, Kerstein & Wang, 2018). Finally, Altay and Ramirez conducted an exploratory study investigating how extreme events affect supply chains in different sectors. They observed more than 3500 disasters, provided by EMDAT database, on over 100 000 companies in a 15 years time span, focusing on leverage, total asset turnover and operational cash flow to proxy firm performance. They found that weather events (namely windstorms and floods) have a positive effect on cash flow while, on the contrary earthquakes do not. They attribute this opposite behaviour to the unpredictable nature of eathquakes that does not allow firms to prepare themselves properly. Moreover, they noticed that flood impact on turnover is dependent on the position that firm occupy in the supply chain: upstream companies enjoy a positive effect, while downstream partners suffer a negative impact (Altay & Ramirez, 2010).

1.3.3 Other articles

In this subsection some of the remaining articles not previously categorized will be discussed. A focus on EWE macro impacts will be favoured, often leading to a cross-sectional overview of the effects over society.

1.3.3.1 Vulnerability, Welfare and Tourism: the Caribbean countries case

Developing countries are often vulnerable subjects to extreme weather events, among them Caribbean countries are periodically impacted by the most devasting meteorological phenomenon, hurricanes. Heinen, Khadan and Strobl studied the link between price effect and potential welfare losses due to tropical cyclones and floods in 15 Caribbean islands. They grouped affected goods in three main categories, “food”, “housing” and “other”, allowing them to observe how individual price effect is small on average, but not negligible at all when grouped. They also find that hurricane absolute damages are lager because they affect richer assets, such as housing and high valued properties, but that the flood impact on food is disproportionately larger for the poor (Heinen, Khadan & Strobl, 2019). Moreover, Hsiang conducted a study on a similar pool of Caribbean countries highlighting the similarity between economic output and labor productivity response to high temperatures. He then separately observed temperature effects and tropical cyclones influence over sectors, modelled simultaneously to avoid their correlation to affect results. The results outline a significative negative effect of temperature on not-agricultural sectors (namely *wholesale*, *retail*, *restaurants and hotels*, *mining*, *utilities* and *other services*) that minimizes the effect on agriculture when comparing also their respective economic relevance in these countries (55% against 10% roughly). On the other hand, cyclones have an overall null effect on total production, suggesting both positive and negative impacts among sectors: while negatively affecting *Agriculture* and *Tourism-related* sectors it also remarkably boosts *Construction* sector, because of its reconstruction role. He concludes raising concern on the persistent effect over time and the disproportional large impact of cyclones on *Tourism-related* sectors, underlining how the effect is mostly driven by the reduction in number of tourists rather than by the decreasing of income per visit (Hsiang, 2010).

1.3.3.2 Climate Change, Adaptation gaps and Urbanization: the Developed countries case

Developed countries are less vulnerable to EWE exposure with respect to developing ones, but that does not mean that they are exempted from casualties and severe consequences. One of the most effective countermeasures, accessible also thanks to larger economic resources, against Climate Change and weather related extreme events is adaptation. Nevertheless, current studies all agree in outlining an adaptation gap that is by now persistent even in richer countries, despite all political and economic efforts in this field. On this issue, Melvin, Larsen, Boehlert et al. raise concern on the rapid onset of climate change effects on Alaska infrastructure. They assess the potential economic damage outcome related to the climate-driven changes in flood, precipitation and permafrost health trends, recommending rapid and substantials efforts in proactive adaptation (Melvin, Larsen, Boehlert et al., 2017).

But Climate Change is not only a risk for Alaska fragile equilibrium, Hsiang, Kopp, Jina et al. in their very comprehensive study estimated climate change influence over US economy, exploring different scenarios. They conducted a cross-sectional analysis over sectors, such as agriculture, crime, coastal storms, energy, human mortality and labour, estimating a loss in gross domestic product of roughly 1.2% per +1°C increase on average temperature. Furtherly, they evidence a disproportionate risk distribution with the result of increasing economic inequality, affecting again the most vulnerable sections of the population (Hsiang, Kopp, Jina et al., 2017).

An important vulnerability of developed country in the context of EWE exposure resides in high-valued concentration of assets and wealth. Obviously, asset exposure is a consistent element in evaluating the damage risk related to an extreme event strike: densely populated cities and concentrated value assets increase this risk. Nevertheless, Grinsted, Ditlevsen and Christensen, in their work, comment how, even if, because increased wealth, the absolute damages of hurricanes have substantially increased in US history, a normalized approach is needed to properly evaluate the trend. Therefore, they “normalized” damages of over a century of US hurricanes to make events comparable, assembling an indicator that considers the equivalent area of total destruction (ATD). Their results are consistent with the absolute trend, suggesting that, even if asset exposure is increased, the frequency of extremely damaging storms is significantly increased also (Grinsted, Ditlevsen & Christensen, 2019).

Finally, human contribution to damages may not only be dependent on population density or value of asset exposed, but also on the health of environment and its correlated capacity of shelter from these events. For instance, Zhang, Villarini, Vecchi et al. focused on how urbanization contributed in exacerbating flood response and total precipitation during hurricane Harvey in 2017. They found that Houston skyline profile and soil sealing severely influenced precipitation and flood, respectively (Zhang, Villarini, Vecchi et al., 2018).

1.3.4 Difference-in-Difference

This subsection is dedicated to Difference-in-Difference methodology. It is the core method adopted in this thesis and this subsection is intended to show how it is also consistently present in the literature. It is adopted in 11 papers of the final pool that represents a 27% of the total, but this share becomes more significant if only economic papers are considered. In fact, among the 26 economic articles 9 of them adopt a Difference-in-Difference method (35%). Furtherly, considering only "Finance" and "Firm Performance" categories, the two most strictly focused on the methodological field of the thesis (micro-founded economic studies), the share increases to exactly 50% (6 out of 12).

The method itself will be described in detail in a further subsection later in this thesis (subsection 2.6.2). Suffice to say that it essentially consists in a method to compare the behaviour of two very similar groups after a particular event affect one of them (e.g. an extreme event that strikes one out of two very similar firms that operate in close conditions). Thus, once defined control and affected groups, the method is capable to highlight outcome differences over time between the affected group and the control, unaffected, one. This is particularly useful in assessing for EWE impacts on outcomes of stricken subjects, because extreme events are clearly located in time and space allowing to sharply distinguish between a before and an after of the phenomenon.

1.3.5 Climate Change influence

This subsection is intended to provide an overview on how the actual scientific literature addresses the relationship between climate change and extreme weather events. To do so a new title-abstarct review on the first pool (850 articles, before binary questions) was conducted with the aim of selecting all papers that concern climate change influence over extreme weather events. From the selected 45 studies, a coherent and consistent narrative appears; even if debate on approaches, forecasts and methods still remains outlining how the scientific interest on this topic is compelling.

Wildfires are a severe and salient phenomenon that tragicly involved wide areas of the world in the recent years (e.g. Californian and Australian fire seasons). The last decade was rich of interest on this topic due to the anomalous and astonishing extent of fire activity. Researchers find evidence of links between this increasing trend in fire activity and the ongoing climate change observing variation in temperature increase, winter snowpack, summer evaporation and precipitation (Holden, Swanson, Luce et al., 2018) (Abatzoglou & Williams, 2016). Some of them do not limit their study to the anthropogenic period (1850-present day) but they go back up to 3000 years ago, retrieving data from sedimentary charcoal, observing how the actual situation is out of the natural climate-fire equilibrium, depicted by the charcoal history, and attributing this divergence to human driven climate change (Marlon, Bartlein & Gavin et al., 2012). Moreover, the co-occurrence of temperature related phenomena, such as hetwaves and droughts, could lead to severe consequences both as wildifire enablers and extreme weather events themselves. On this topic researches have no doubt in stating how heatwaves in recent decades have been already affected by human forcing climate effects (Hansen, Sato & Ruedy, 2012) and raise concern on the alarming projections of more intense, frequent and long lasting phenomena by the end of the century (Meehl & Tebaldi, 2004). Other studies address drought focusing on trend projections in localized areas, such as the Amazon and western US, concluding that climate change have already had a role in the last decades drought anomalies and that it is going to be a key driver in future trend forecasts (Duffy, Brando, Asner & Field, 2015) (Diffenbaugh, Swain, Touma & Lubchenco, 2015). Another US study highlights how trends of concurrent extreme events, namely heatwaves and droughts, have already increased in the period 1990-2010 with respect to the baseline level (1960-1980) (Mazdiyasni & AghaKouchak, 2015), while another, in contrast, debates how PDSI (Palmer Drought Severity Index), the most common used drought indicator, could produce misleading results due to biases related to its simplicity when used in the contest of climate change (Sheffield, Wood & Roderick, 2012).

Anyway, all of them substantially agree on the influence of climate change on future droughts severity and frequency.

Storms are a phenomenon in depth studied within the literature. In particular, attention focuses on tropical cyclones, even if also thunderstorms are considered (Trapp, Diffenbaugh, Brooks et al., 2007), because of their severe impact on lives and assets. Several papers find evidence of the relationship between climate change, through radiative forcing anomalies, and TC response (Trouet, Harley & Domínguez-Delmás, 2016) (Grinsted, Moore & Jevrejeva, 2012), while others stressed on how climate change particularly affects the frequency of very extreme and intense events rather than their overall frequency (Bender, Knutson, Tuleya et al., 2010). Moreover, TCs are also connected to flood risk. In fact, global warming may induce TC slowdown in turn related to more intense precipitation and water accumulation increasing significantly flood risk (Lai, Li, Gu et al., 2020). Furtherly, coastal flooding is enhanced by concurrent TC and increasing sea levels; in particular surging seas seem to be the most relevant forcing in the increasing trend of coastal flooding risk during anthropogenic era, with respect to baseline levels of pre-industrial period, (Reed, Mann, Emanuel et al., 2015) and also for future forecasts of the century (Garner Mann, Emanuel et al., 2017). On the other hand, river floods are influenced by the changing climate mainly through variation of precipitation accumulation, with projected scenarios that shows higher frequency of very intense events (Neelin, Sahany, Stechmann & Bernstein, 2017). In this case, the response is not univocal, such as increasing sea levels, but differentiated throughout places around the world (Dankers, Arnell, Clark et al., 2014). Focusing on Europe, river floods are projected to increase particularly in northwestern area, due to increasing precipitation, while presenting an inflection in southern, due to reduced rainfall, and eastern areas, related to decreasing snowpack (Blöschl, Hall, Viglione et al., 2019). A significant and documented example of the consequences of changing climate on flood risk in Europe is provided by the England and Wales flood in autumn 2000 where anthropogenic forcing was found to have severely increased the probability of flood occurrence (Pall, Aina, Stone et al., 2011).

1.4 Research question

This section is intendend to define the research question of this thesis. From the summary and some of the features of the literature previously outlined, it seems clear that further research is possible and necessary in several directions.

At least three main aspects from the literature concur in defining premises for the research question.

Firstly, the large diffusion of the Difference-in-Difference methodology, as previously discussed, can be interpreted as an evidence of scientific community approval of the effectiveness of this method in the field of EWE driven shocks on economic activity.

Secondly, geographical distribution of countries covered by studies returns a clear situation. Most of the studies addresses US territory, mainly focusing on reactions of banks and managers to exogenous shocks, while almost the other half refers to developing countries' exposure and consequences to weather events. Only two papers directly issue Europe and only one of them is a micro-founded study (*Borrowers under water! Rare disasters, regional banks, and recovery lending*, Koetter, Noth & Rehbein, 2020). Thus, considering the opportunity of detailed data availability on both extreme events and financial information in developed countries, research developments on Europe geographical area are open and compelling.

Finally, among the wide list of articles that present a micro-founded study on EWE impacts, only three of them address data at a company level. Furtherly, two of them perform a cross-sectional comparison over sectors and all of them refers to financial information of large public companies or publicly traded firms. This seems reasonable when considering the data ease of access, but on the other hand small-sized firms are those that may be more exposed to the occurrence of an extreme event.

For all these reasons, it is possible to identify an unexplored literature aspect, that defines the research scope of this thesis: *a multisectoral analysis of extreme weather event impacts* *on firm performance in the European area, founded on company level financial information and with a focus on SMEs.*

The statement of the research question closes this chapter, with the aim of offering a novel and significant contribution to the existing literature on EWE shocks on economy.

Chapter 2 Methodology

The following Chapter is going to illustrate the methodology adopted in this thesis. The steps that led to the collection and assembly of complete datasets are described and an explanation and discussion of the model employed is provided.

2.1 Data sources

In order to assess the relation between firm performance and extreme weather events, information have been retrieved from both financial and weather based datasets at European level. Additional sources have also been necessary for integration and they will be discussed later.

2.1.1 Financial database source

The designed choice for the financial database was the Orbis[3] dataset, from Bureau van Dijk.

Access to the database was possible thanks to the license provided by the Politecnico di Milano to the research group. The login to the platform was allowed exclusively through the university internal network, for this reason a proxy connection was implemented to make operations on sources and, in particular, the export procedure easier.

Orbis is a very complete database that reports information for more than 375 million firms all over the world. It provides several different kinds of information related to companies like unique identification code (BvD ID), location (with some critical aspects better discussed later), activity and industry, financials, directors, managers and advisors, ownership, stock and earnings estimates, credit default swaps, M&A, intellectual properties and others. As a logged user it is possible to set queries to filter the desired information, then the results can be downloaded in several formats respecting the most limiting factor concerning the size of the single file to be exported. Lastly, all the information is issued in a standard form that allow the comparison of firms independently from country origin. For these reasons Orbis is one of the most recognized and used source for micro-founded economic studies.

[3] https://orbis.bvdinfo.com

The version of Orbis accessible to the research group was the ordinary version that limits the number of data to be examined at a maximum of 10 years before the last available year (2018 for this study). Manly because of this, as well as for management concerns of bigger datasets, the company information is limited to the period from 2009 to 2018.

2.1.2 Extreme Weather Events source

The information for the weather dataset was taken from EMDAT[4], the Emergency Events Database launched in 1988 by the Center for Research on the Epidemiology of Disasters (CRED) with the initial support of the World Health Organization (WHO) and the Belgian Government. The dataset is compiled from varius sources (UN agencies, non-governmental organizations, insurance companies, research institutes and press agencies) and it contains data on more than 22 000 disasters world spread from 1900 to present day.

EMDAT is one of the main sources of information about all types of natural hazards. It provides a bunch of information such as unique disaster identification number, event type and sub-type, geographical and temporal information, physical characteristics of the disaster, status of intervention, sources of information, human and economic impact. It embraces different types of events that classifies in two macro groups (Natural and Technological); the latter refers to such industrial accidents of varius nature from chemical spilling to explosions and gas leakages; the natural events instead further differentiate into geophysical, meteorological, hydrological, climatological, biological and extra-terrestrial sub-type. For the needs of this thesis only European natural events of meteorological, hydrological and climatological sub-types are considered because of their strict relation to the weather and to climate change. The extraction of data was also limited to the period from 2008 to 2018 to properly match with the company information available.

The most common alterative to collect data on different kinds of extreme events is NatCatSERVICE of Munich RE[5]. It uses different criteria to include extreme events within

[4] https://public.emdat.be
[5] https://www.munichre.com/en/solutions/for-industry-clients/natcatservice.html

the database, in particular it focuses more on insured and overall losses. In NatCatSERVICE small-scale property damage and/or one fatality are sufficient to be included as extreme event, on the other hand EMDAT needs at least 10 or more casualties and/or 100 or more people affected and/or declaration of a state of emergency and/or call for international assistance. The stricter criteria and the public nature of EMDAT makes it the most feasible choice for this thesis.

2.1.3 Additional sources

Auxiliary data sources were needed to integrate missing information in the datasets. For this reason also Eurostat and Bing Maps Location API were exploited as useful data integrators.

Eurostat[6] is the statistical office of the European Union, built to collect comparative statistics and data on Europe. It works in partnership with National Statistical Institutes and other national authorities in the EU Member States forming the European Statistical System (ESS). It includes also statistical authorities of countries that belong to the European Economic Area (EES) and Switzerland. In the scope of this thesis the Geographic Information System of the Commision (GISCO) granted useful information on NUTS3s classification (e.g. GIS shapefiles). The NUTS (Nomenclature of Territorial Units for Statistics) classification is the key geographical information to match financial and event datasets.

The Bing Maps REST Services Application Programming Interface (API)[7] through a Representational State Transfer (REST) interface allows to perform different operations such as creating static maps with pushpins, geocoding an address, retrieving imagery metadata or creating a route. It is a free to use tool within limits on the maximum number of operations allowed for without a paid subscription, it was particularly useful in recover missing address information in both datasets.

[6] https://ec.europa.eu/eurostat/home
[7] https://docs.microsoft.com/en-us/bingmaps/rest-services/locations/

2.2 Dataset construction

The steps that led both datasets to be assembled are reported in the following sections. The two databases were processed separately till they where complete and ready to be used during the analysis. A final merge into a unique dataset was not useful because of the nature of the analytical procedure: each event is studied independently. Thus, the merging of datasets was performed during the analytical procedure (Section 2.6).

2.2.1 Economic dataset

Collection of data for this database was a relevant step through the assembling process, then appending, merging and raw cleaning were performed.

2.2.1.1 Data collection

As mentioned above, financial information on firms was collected from Orbis platform. It allows to fill and save a query with custom parameters. The query filters requested data on the base of inserted parameters and shows results of the search.

The main collecting criteria was to download the wider group of European companies as possible in order to have a reliable dataset also for further application and studies within the research group.

The first bulk query asked for the following parameters to be respected:

- *Region:* Western Europe, Easter Europe (All European territory).
- *Completeness:* filtering companies with at least one report of operating revenues (turnover) available in the time period selected.
- *Time Period:* all allowed available years (from 2009 to 2018).
- *Firm category:* Public authorities, Governments and States were excluded.

To avoid the display of not relevant information out of the search also a custom “view” was designed. The “view” option is an Orbis tool that allows to filter and actually display in tables only the desired information about a company. The view parameters are listed, the following list is the result of several integrations implemented work in progress and guided by the evolving needs of the research group:

- *Company name:* complete name of the firm.
- *ID number:* BvD ID code, a unique identification code assigned by Orbis to each company.
- *Contact information:* coordinates, address, postcode, city and country.
- *Legal & account information:* size classification, status, status date and date of incorporation.
- *Ownership data:* number of subsidiaries and number of branches.
- *Industry classification:* NACE Rev. 2, core code (4 digits).
- *Balance sheet:* fixed assets, intangible fixed assets, tangible fixed assets, other fixed assets, current assets, stock, debtors, other current assets, cash & cash equivalent, total assets, shareholders' funds, capital, other shareholders' funds, non-current liabilities, long term debt other non-current liabilities, provisions, current liabilities, loans, creditors, other current liabilities, total shareholders' funds and liabilities, working capital, net current assets, enterprise value, number of employees.
- *Profit & loss account:* operating revenue (turnover), sales, costs of goods sold, gross profit, other operating expenses, operating P/L (EBIT), financial P/L, financial revenue, financial expenses, P/L before tax, taxation, P/L after tax, extr. and other P/L, extr. and other revenue, extr. and other expenses, P/L for period (net income), export revenue, material costs, costs of employees, depreciation & amortization, other operating items, interest paid, research & development expenses, cash flow, added value, EBITDA.

The query result showed a table for more than 22 million firms (one for each row) and more than 500 columns mostly composed by the 50 time-depending financial variables. All this information needed to be extracted from the platform into a practical format like Excel. The procedure was complicated by Orbis limitations regarding the maximum number of companies (rows) to be exported at a time. This limit stands at 1901 companies per export (.XLSX format) and it prevents the dataset to be downloaded all at once.

Obviously, the whole process could have been carried out manually, but it would have taken far more time and it would have been more subject to human errors, as well as very punishing for operators.

For these reasons a couple of Python scripts were implemented to carry out the extraction process. Both with different tasks: the first implementing browser automation to fill the export section of Orbis platform from where the files could be downloaded; the second designed to automating the download process when the exports were all ready to be extracted. The scripts were based on Selenium WebDriver[8] tool, a collection of open source APIs used to automate web applications. This Python additional component was extremely useful in implementing the browser automation.

The need for two separate scripts was subjected to material time delay from the input of the first script and the moment when the exports were actually ready to be downloaded. The high number of requests that Orbis was asked to process at once often overloaded the platform leading to the need of at least a 24 hours delay before running the second script. This overload hindered the procedure also with the occurrence of several bugs that corrupted some exports impeding their actual download. This specific problem was overcome by an integration control, further explained in detail.

The overall raw Excel files download took approximately 15 weeks and ended up with more than 11 000 files downloaded.

2.2.1.2 Data appending

The appending procedure, or vertical concatenations, is described by the combining of two or more different datasets of the same nature (same kind and number of columns) resulting into a single larger database with all the observations (rows).

All the 1900-firms export files needed to be appended into a single database. Dealing with such a large size of information could be very tough. Thus, the appending process was carried out exploiting the versatility of Python in dealing with dataset of such a size.

[8] https://www.selenium.dev/

For this specific task another Python tool showed up to be extremely useful: Pandas[9] module. Pandas is a very powerful and flexible data analysis/manipulation tool for Python that grants several ready to use functions for the purposes of this thesis.

A first vertical concatenation was performed appending 1000 Excel export files at once into an approximately 1 900 000 firms file. The resulting blocks were saved both into .CSV and .DTA formats. The first to go across further Python steps, the latter to be usable via STATA. These subsequent appending steps were a trade-off between aggregation (the final goal) and manageability.

Agility in handling the information was fundamental to optimize integration and preliminary cleaning processes were needed before the final append into a single .DTA file. This final append was the last time Python was used for the purposes of this thesis, establishing the definitive adoption of STATA as reference working software.

STATA[10] is a powerful and widespread software for data analysis, statistics, visualization and managing of large datasets. Licences for the STATA 16/SE (64-bit) version was provided by Politecnico di Milano to the research group.

2.2.1.3 Data integration

Data integration, as explained above, was made necessary due to the evolving needs of the research group and to correct bugs in the export process. This section describes what kind of information was added and how the procedure took place.

The first integration concerned the missing companies due to export bugs. To overcome this deficiency the full list of BvD IDs was downloaded with a specific query. The information requested was far lighter than the first massive query and this characteristic was enough to avoid bugs in the export process. Then the full list of BvD IDs was compared with the companies in blocks. BvD ID is the unique identification number for companies in the platform so it is the only appropriate and independent indicator for this kind of search. All BvD IDs with no match were used as filter to the next query in order to retrieve the missing

[9] https://pandas.pydata.org/
[10] https://www.stata.com/

company data. The new downloaded information was then appended into blocks as explained above.

The second integration regarded all additional general information not subjected to the first query. Specifically, *Contact information, Legal & account information, Ownership data, Industry classification* were integrated.

As before, the new downloaded data was relatively light with respect to the blocks and was grouped into a single file and handled via Python. Exploiting the merging functions of Pandas, each 1 900 000 firms block was compared to the additional file (containing the new information and the BvD IDs). BvD ID was used as key in the merge: when a BvD ID in both files matched, the new information was added to the considered block. At the end of the process, after all blocks were compared, the number of rows (companies) was unchanged but with all the new information added as new columns.

After this integration step it was noticed that location information, one of the fundamental aspects for the final matching between firms and weather events, was reported with very differentiated accuracy. After a preliminary check only a little less than 1 800 000 companies reported longitude and latitude information. To avoid a massive loss of available data, a more spread geographical indicator was selected. The most appropriate trade-off between accuracy and completeness in the dataset was the postcode information. It was therefore used as leading criterion for the retrieving of information for all companies with missing coordinates from Orbis.

A Python script was implemented exploiting the GeoPy[11] extension together with the Bing Maps Location API. Postcode, city and country information were extracted and appended into a single string and then subjected to the Bing Maps Location API to retrieve coordinates for companies that lacked them.

[11] https://geopy.readthedocs.io/en/stable/

This procedure led to the retrieving of more than 8 500 000 company coordinates. Then this information was integrated into blocks with the same process described above in the subsection.

The last data integration needed was area and population of each NUTS3 region at which the observed company belong. The adoption of NUTS region classification will be discussed in the following sections, but it is mentioned here as the key factor for merging this new additional information to the company blocks.

2.2.1.4 Preliminary cleaning

With the word "preliminary" are intended all cleaning operations performed on the company dataset before the final cleanings, carried out on the final firms/events datasets.

The first cleaning consisted in the dropping of all absolute duplicate observations (rows): each time two complete rows reported the exact same information, one of them was discarded. This procedure was firstly conducted on the company blocks with Pandas functions via Python, then again on the complete firm dataset in STATA.

The second cleaning affected almost blank lines in the company blocks. It was probably a bug in the Orbis exports and presented the duplication of BvD ID and company name, but the total absence of other information in the row. Again with Pandas tools, rows with this anomaly were detected and dropped.

The last preliminary cleaning was intended to dig deeper into the dataset allowing to consider only companies that represent the European firm sample properly. In particular the presence of other financial variable, out of operating revenues (turnover) considered in the original query, was taken into account. The new dropping criteria was subject to the total absence of *Total assets* and *Sales* information during the whole time period considered (2009 – 2018).

All the preliminary cleaning finally led to a sample of approximately 13 million companies (one for each row).

2.2.1.5 Financial variables definition

This section describes the process that led to the final shape of financial variables. The choice and the design of variables considered several aspects.

Firstly, the choice of variables fell on classical and immediate firm performance indicators such as Total assets, Operating revenues and EBITDA. Starting from these three indicators, Cost of Operations (Operation revenues – EBITDA) and Return over Assets (EBITDA/Total assets) were built. In particular, Cost of Operations (CO) was designed to compensate the scarcity of availability of other cost variables already present in the dataset.

Furtherly, in order to limit the positive skewed distribution effect on data, logarithmic values of variables were considered. Obviously, to avoid loss of information from null values, a unit value was added to the argument of the logarithm. The only exception was ROA, in fact it is the only variable that is not positively skewed, and rather it was applied symmetric winsorization at 1% level (a method that consists in substituting extreme figures to the percentiles of 0.1 and 0.99).

After these considerations, it is possible to list the four financial variables considered in this thesis:

- *Total assets (TA):* logarithmic value of total assets. Specified by company and by year (2010-2018), dimensionless.
- *Operating revenues (OR):* logarithmic value of operating revenues. Specified by company and by year (2010-2018), dimensionless.
- *Cost of operations (CO):* logarithmic value of the difference between operating revenues and EBITDA. Specified by company and by year (2010-2018), dimensionless.
- *Return over assets (ROA):* winsorized value of EBITDA over total assets. Specified by company and by year (2010-2018), dimensionless.

2.2.2 Extreme Weather Events dataset

Differently from the economic dataset, the extreme weather event database major issue was related to transformation of data, rather than collection, in order to make it suitable for the scope of the research.

2.2.2.1 Data collection

The event database was downloaded from the EMDAT database website. It presents itself as an Excel file accessible after registration to the website as academic organization.

2.2.2.2 Data processing

The original Excel file incurred then several transformations. The first was a massive cut to keep the focus on European climate related events. Because of this all events not belonging to geographical Europe and meteorological, hydrological and climatological sub-types of natural events were discarded. Then also the time period was restricted considering years from 2008 to 2018. Keeping also 2008 year was a choice dictated by caution and data completeness without implication of significant downsides.

As listed above, EMDAT database reports several useful information about events, but one of the most relevant for this thesis (location) is stored in an unsuitable form for the uses of the research. In particular if an event interests more than one single location (e.g. countries, provinces and districts) they are anyway all reported in the same cell.

This issue was handled exploiting the Excel feature to separate cell content into different rows using separators as reference (e.g. comma, semicolon or space). This passage was performed manually and without auxiliaries of macros, because of the very heterogeneous nature of report and listing criteria: different separators among countries in listing locations, also with exceptions within the same country.

Amid this process, that took approximately four weeks, the geographic key criterion for the data matching between financial and event dataset was decided: NUTS regional classification, as mentioned above.

Nomenclature of Territorial Units of Statistics (NUTS)[12] is a hierarchical location framework for dividing up the EU and the UK economic territory for several purposes: socio-economic analysis of regions at different geographical levels, EU regional policies framing and collection, development and harmonization of European regional statistics. The NUTS classification is managed within Eurostat by the Geographic Information System of the Commission (GISCO)[13] and it is structured on four different levels (NUTS0, NUTS1, NUTS2 and NUTS3). The most granular level is NUTS3 and for this reason it was chosen as location reference by the research group. Other levels represent progressively larger surface areas following different criteria depending on countries. Generally, NUTS0 is associated to the whole country surface, NUTS1 represents macro-regions, NUTS2 mainly describes administrative regions and lastly NUTS3 corresponds to provinces/counties. As mentioned before, this division is indicative and each country, depending on several factors (e.g. internal administrative organization, population, surface extension), applies the classification with slight differences.

This classification does not include all the countries within geographical Europe, so all events belonging to countries not included in the classification were discarded (Belarus, Bosnia and Herzegovina, Kosovo, Moldova, Russian Federation and Ukraine). Through these procedures (cell separation and country drop) the original 447 observations were transformed into a 2396 rows database.

The following step consisted in the collection of geographical coordinates information of each observation. This step was needed to prepare the dataset for the next step to be performed with QGIS software. The goal was to finally reach a NUTS classification for

[12] https://ec.europa.eu/eurostat/web/regions-and-cities/overview
[13] https://ec.europa.eu/eurostat/web/gisco

events location. This step was performed through several sub-processes. The first involved the Excel Geography Data Tool.

The Excel Geo-Tool can be applied to cells containing locations, then if it finds a match it gives back interesting information (e.g. coordinates, country, region, population, area, regional identification). Therefore, all locations corresponding to cities could by easily discriminated from other location levels because of the presence of coordinates information.

All the remaining locations belonged to different geographic levels (e.g. country, region, province), thus further discrimination criteria were needed. It was then noticed that all matched locations corresponding to states did not reported the Excel Geo-Tool information related to country (e.g. location “Milan” reported “Italy” under country information, but the same information for location “Italy” was systematically missing). This peculiarity was very useful to distinguish this category from others.

Therefore, a preliminary classification was performed. They were established four different categories (named Dimension by the author) and they were coded as following:1-Cities, 2-Province, 3-Country and 4-NotFound. The latter dimension corresponds to all locations that Excel Geo-Tool had not matched, while first and third dimensions have been described above. The remaining dimension finally (2-Province) deserves further explanation: it contains all locations not belonging to other dimensions, but they are not limited strictly to provinces. In fact, depending on the country considered, several different administrative entities could be contained in this dimension (macro-regions, sub-regions, provinces, counties).

Before further proceeding with 2-Province dimension, a couple of checks were carried out. The first involved 4-Not-Found locations, that were manually examined and corrected where possible. Then also a “Country Check” was executed: country information already present in the original EMDAT database was compared with the one founded with Excel Geo-Tool. All mismatches were manually checked and corrected where possible, otherwise added to 4-Not-Found dimension and then discarded for a total amount of 145 observations.

Finally, the 2-Province dimension separation was handled. As mentioned above, the main issue of the entire process was related to the dimension separation: Macro-Regions like “North Italy”, actual regions like “Piemonte” and provinces like “Provincia di Alessandria” fell all in the same dimension indiscriminately. Each country uses peculiar criteria and is organized into different kind of entities (e.g. Regions, Macro-regions, Provinces, Counties) so, taking the NUTS level criterion as reference, each country administrative regional organization was manually searched on the net.

For all European countries involved NUTS levels 1, 2 or 3 have been associated. Recalling the previous example, for instance NUTS1 for “North Italy”, NUTS2 for “Piemonte” and NUTS3 for “Provincia di Alessandria” were respectively coupled. In doing so, useful information from Excel Geo-Tool, defined “Abbreviation”, helped simplifying the process. It is a code system used to identify geographical administrative entities. It proved to be very useful because it provided an unambiguous indicator of administrative level within countries. Taking Italy as an example, the “Abbreviation” code was reported only by the administrative level “Regione”, hence it helped to immediately distinguish from the other fundamental administrative level of interest “Provincia”. Unfortunately, this indicator, even within the same countries, was not exempted from exceptions, that hindered and slowed the process.

NUTS0 level was then matched to the 3-Country dimension previously defined. Using the NUTS criterion, all observations were futherly separated in four new sheets: NUTS0, NUTS1, NUTS2, NUTS3 and Cities. Finally separated and classified, most of the locations, with the only exception of 1-Cities dimension, still lacked in coordinates information.

To gain easily access to coordinates, the “Major City” information of the Excel Geo-Tool was exploited. Among several others, the Excel Geo-Tool also give access to information about capital and major city. The major city not always corresponds to the administrative capital, but it was preferred as geographical indicator. The choice was made on the base of the fact that it was less frequently missing in the dataset with respect to capital city and that for our purposes both indicators would have led to the exact same result in classify the NUTS level.

Therefore, all major cities were in turn processed with the Excel Geo-Tool returning the coordinates information, generally present for cities. As well as previous steps, all exceptions were manually checked and corrected where possible. In particular, the Bing Maps Location API was again used to retrieve lacking coordinates when the Excel Geo-Tool failed in providing them.

Finally, the five sheets (NUTS0, NUTS1, NUTS2, NUTS3 and Cities) were exported as separated .CSV files. Even if each observation is now associated to a NUTS classification level, the actual NUTS identification code is still missing. Hence, further processing via QGIS software was still needed.

2.3 QGIS shapefile matching

QGIS[14] is a free and open source software part of the Open Source Geospatial Foundation (OSGeo) that allows to perform several operations on geospatial information, such as create, edit, visualize, analyse and publish. In the purposes of the thesis, it allowed to match NUTS ID codes, reported into shapefiles, with event coordinates of the .CSV files arranged before.

Firstly, .SHP shapefiles were downloaded from Eurostat. The latest version of the document at the time of execution was NUTS 2016 version (14/03/2019). Once imported into QGIS, they draw several map layers corresponding to each of the four NUTS levels.

NUTS classification was elected as the reference location framework for both financial and event observations. As anticipated above, coordinates referring to companies were estimated from postal codes reported in the Orbis database. Considering exact company position would have made little sense taking into account Orbis constraints: coordinates were mostly retrieved from postcodes rather than actual addresses and the reported position of registered offices refers to headquarters that could not necessarily correspond to actual production

[14] https://qgis.org/en/site/

facilities. Thus, for the latter consideration specifically, an assumption is introduced: production facilities or offices are assumed to being placed close to registered headquarters. Similar considerations can be made for the EWE dataset: weather events can hardly be exhaustively represented by a specific point on a map.

Anyway, because of these imposed approximations, the choice for a location criterion fell on NUTS3 (province/county scale, officially defined "small regions for specific diagnoses") as a structured, spread and recognized geographical classification.

2.3.1 Economic Dataset

Registered office positions in the form of coordinates were exported as a .CSV file from the assembled dataset. The .CSV file was then imported into QGIS as a layer.

Each observation presented longitude and latitude information, so in the process of creating the new layer QGIS assigned a point on the map on the base of coordinates supplied. This new layer containing all company positions was then compared with the NUTS3 layer of the shapefile. The "Join attributes by location" feature was used to carry out the process. This feature returns a new layer consisting of the company position layer integrated adding a column which reports NUTS3 ID codes.

The new layer was then exported into a .CSV file It actually represented a coordinates/NUTS3 conversion file and it was then used to integrate the original Economic dataset via STATA, exploiting coordinates as key.

Finally, the last cleaning step on the financial database was conducted. It involved all observations that lacked NUTS information, that were dropped. This batch represents all countries excluded by GISCO classification (e.g. above mentioned in Section 2.2.2) and all companies of which coordinates could not be retrieved.

After this cleaning, the complete dataset finally counts 8 655 009 observations, each representing a unique company.

2.3.2 Extreme Weather Events Dataset

Also for event observations, ID codes were retrieved via QGIS software. However, this time Eurostat provided NUTS shapefiles were not enough. Hence, a NUTS_CHAIN organization was assembled specifically for this purpose.

Because NUTS3 classification was elected as the reference for the merging between financial and event information, observations belonging to other levels had to be somehow converted. It was decided to assemble a NUTS_CHAIN key capable to go back to the corresponding NUTS3 levels within the selected one. For instance, if an event occurred in a NUTS2 level, the NUTS_CHAIN provides information on all NUTS3 levels contained within the considered NUTS2 level. In other words, the NUTS_CHAIN contains all the ID codes of each NUTS3 level and the associated ID codes of all the NUTS levels above it (in which the NUTS3 is included).

The NUTS_CHAIN was assembled starting from Eurostat shapefiles and exploiting the "Join attributes by location" QGIS function. Therefore, using that feature, a chain of levels going back till NUTS0 was associated to all NUTS3 levels (e.g. NUTS3 "Provincia di Alessandria" →NUTS2 "Piemonte" → NUTS1 "Nord-Ovest" → NUTS0 "Italia").

All .CSV files were then imported into QGIS as separate layers. Creating the new layer, as well as for company locations, QGIS assigned a point on the map on the base of coordinates supplied. Each layer was then compared with the corresponding NUTS level shapefile, exploiting again the "Join attributes by location" feature, to associate the ID code to each observation.

Finally, all the modified layers were integrated with NUTS_CHAIN information. Observations in layers above NUTS3 level (NUTS0, NUTS1, NUTS2) were duplicated in order to report information on all the NUTS regions contained within. Observation belonging to NUTS3 and Cities layers, obviously, did not needed duplication, but they were just integrated with NUTS_CHAIN information about above NUTS levels.

All layers, finally sorted and organized in a common structure, were exported and merged into a single .CSV file and then immediately imported into STATA. Then, a couple of final checks were performed. The first involved the drop of absolute duplicates (exactly alike rows); the second was a country check that compared the country name provided by EMDAT and the new information added with NUTS integration, all not matching observations were dropped. The amount of dropped rows after these checks was negligible with respect to the total number of observations.

2.4 Final operations

At this point both datasets are almost in their final form. Anyway, because of the nature of the analytical method used, the two datasets are not merged into a unique file. Potentially the same NUTS3 could have been affected by different events, as it is likely to happen, so merging the datasets is not possible without overlapping. For these reasons, during the analysis, each event is extracted from the event database and singularly merged with the financial dataset. At the end of the analysis, before merging the next event, the financial dataset is restored to its original form.

As part of the analytical process, the databases are merged together via STATA through a .DO file, a STATA feature comparable to a script, used to run the analysis. The merging process involved a master/slave approach using the economic dataset as master, to avoid loss of financial information, and NUTS3 codes as key.

Another important consideration that needs to be done regards the form of the financial database. It is a wide-shaped database in its pre-analysis final form: each row represents a unique company, while the time-variant characterization of variables is given by subsequent columns (e.g. Total assets of company X in the period are reported all in the same row in adjacent columns, one by year). The wide-shaped dataset is used in the first part of the analysis, namely the control group matching, while in the second part, namely the DID estimation, the dataset is needed in a panel form, long-shaped. The long-shape represents the time-variant characterization of variables duplicating the time-invariant information in rows for each year (e.g. Total assets of company X in the period are reported in adjacent rows, one by year, all in the same column). For this reason, a database reshape is needed during the analytical process, from wide to long, without any actual information loss.

2.5 Descriptive Analysis

This section has the aim of providing an overview of company and climate datasets statistical potential and characteristics. Some of these characteristics has guided the author in the choices during the model drawing phase.

2.5.1 Financial data description

As represented in the following table, the financial dataset covers a wide range of countries in the European geographical area.

Count of firms by country

Country	Count	Country	Count
Albany	178	Lithuania	29.278
Austria	40.924	Luxembourg	12.083
Belgium	31.190	Malta	2.847
Bulgaria	493.902	Montenegro	11.544
Croatia	134.247	Netherlands	21.513
Cyprus	987	Macedonia	82.937
Czech Republic	151.894	Norway	250.766
Denmark	37.703	Poland	215.178
Estonia	131.018	Portugal	365.898
Finland	179.128	Romania	739.165
France	716.451	Serbia	104.884
Germany	230.759	Slovakia	166.239
Greece	23.165	Slovenia	128.869
Hungary	391.407	Spain	676.283
Iceland	30.789	Sweden	154.893
Ireland	30.482	Switzerland	355
Italy	752.744	Turkey	16.078
Latvia	124.372	United Kingdom	407.568
Liechtenstein	23		
		Total	**6.887.741**

Table 2.1 - Count of firms by country in the complete database

Here are reported also countries later excluded from the study because not contained in the EWE database (i.e. Cyprus, Finland, Iceland, Liechtenstein, Malta and Turkey). Moreover, with few exceptions (i.e. Albany, Cypus, Liechtenstein, Switzerland), all countries adopting a NUTS classification report at least a thousand companies, with a maximum represented by Italy.

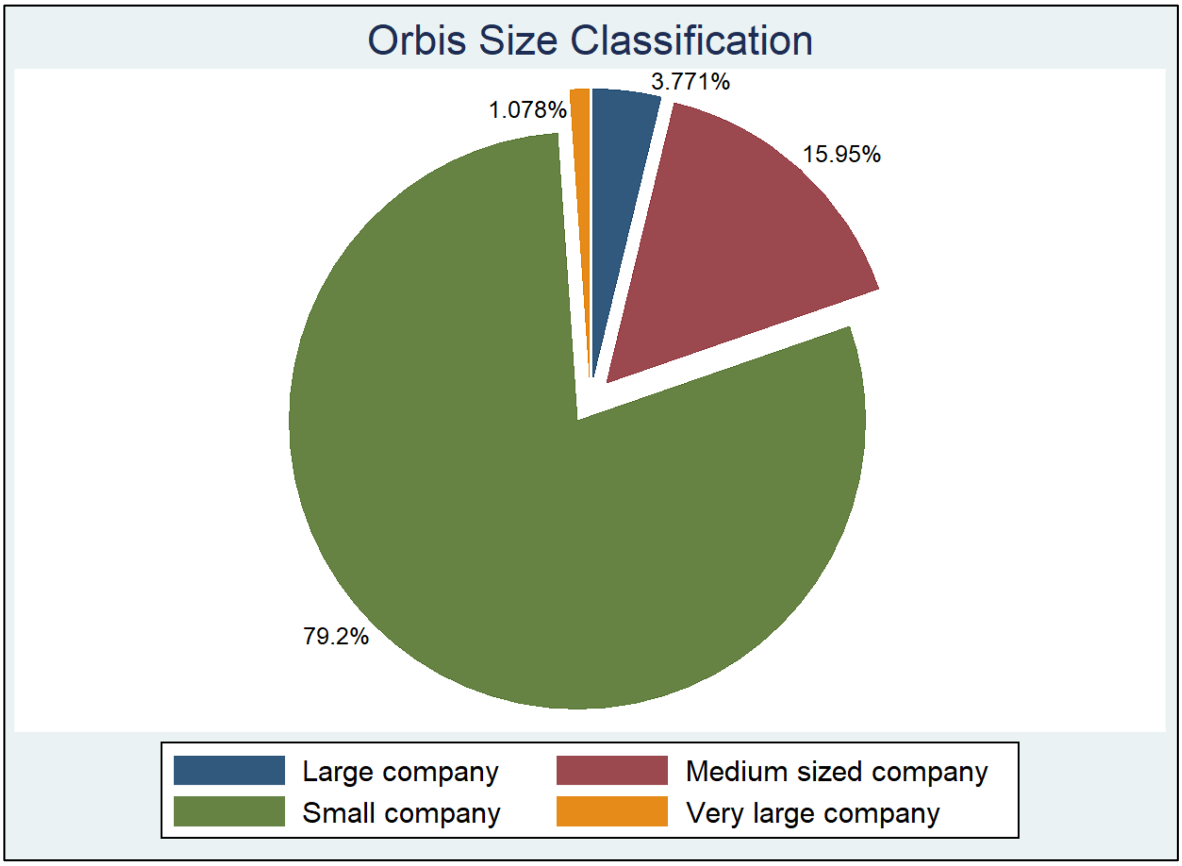

Figure 2.1 - Time invariant size classification, provided by Orbis.

Another interesting consideration is related to SMEs distribution in the dataset. As mentioned before, this work focuses on them because of their higher vulnerability to extreme weather events (Runyan, 2006) but also because of their representativity in European companies distribution, as shown in the chart. Orbis provides its own size classification for companies: very large if it either has 100+M€ operating revenues, 200+M€ total assets, 1000+ employees or if it is publicly listed; large if it has at least 10M€ operating revenues, 20M€ total assets or 150 employees; medium sized if it has at least 1M€ operating revenues, 2M€ total assets or 15 employees; small if all the previous condition are not satisfied.

Anyway, this size classification criterion, as it is reported, is time invariant. In fact, Orbis provides a single size classification based on the last report available. In the scope of this thesis, a time invariant classification is not enough because of the time variant nature of a difference-in-difference analysis: a large sized company affected by an extreme event could worsen its performance enough to downgrade to a lower size level, while an unaffected SME could benefit from the new market situation enough to upgrade to an upper level.

For this reason, an independent size classification is performed for each year reported as a time variant variable. In particular, the Eurostat[15] definition is used: a SME is a company with less than 50M€ operating revenues, less than 43M€ total assets and less than 250 employees. This new variable will be successively mentioned during the model definition in next section.

Furtherly, completeness of variables in the dataset is addressed. The next graph represents the count of non-missing observations in the dataset during the time span considered.

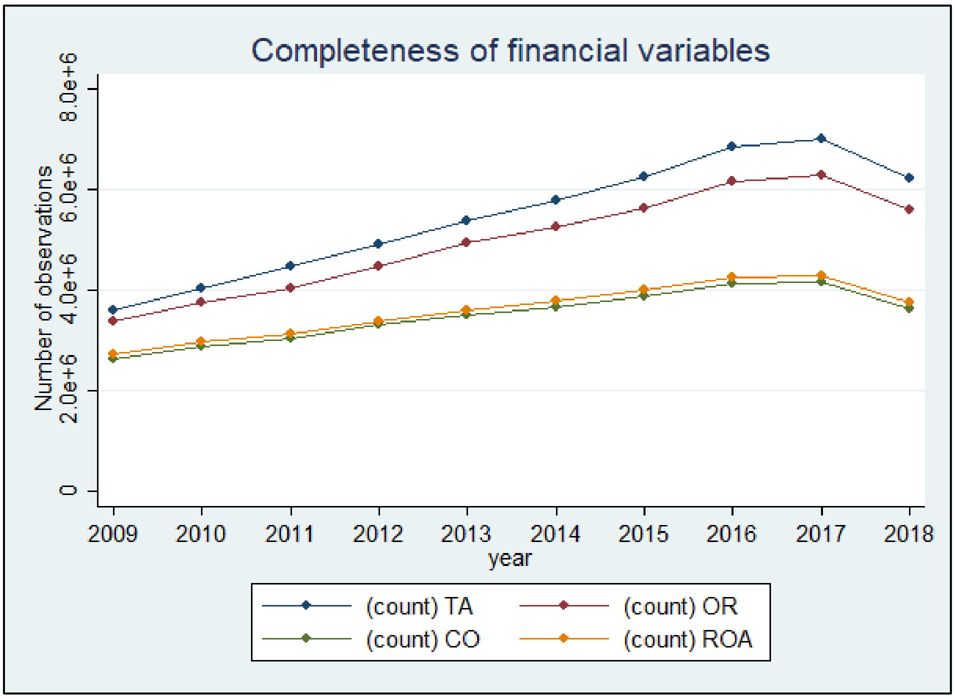

Figure 2.2 – Count of financial variables observations reported in the dataset.

[15] https://ec.europa.eu/growth/smes/sme-definition_en

The graph shows an increasing trend over time with its maximum in 2017. TA and OR are the most represented variables, while CO and ROA, even if abundant, are slightly less present in the dataset. Furtherly, CO and ROA show a very similar trend, this is due to the fact that they are both built starting from EBITDA combined with another financial variable, Operating revenue and Total assets respectively. Among these variables, EBITDA acts as the limiting factor, returning a lower availability for CO and ROA with respect to TA and OR.

Moreover, the maximum availability in 2017 is coherent with the company age distribution shown in the following graph.

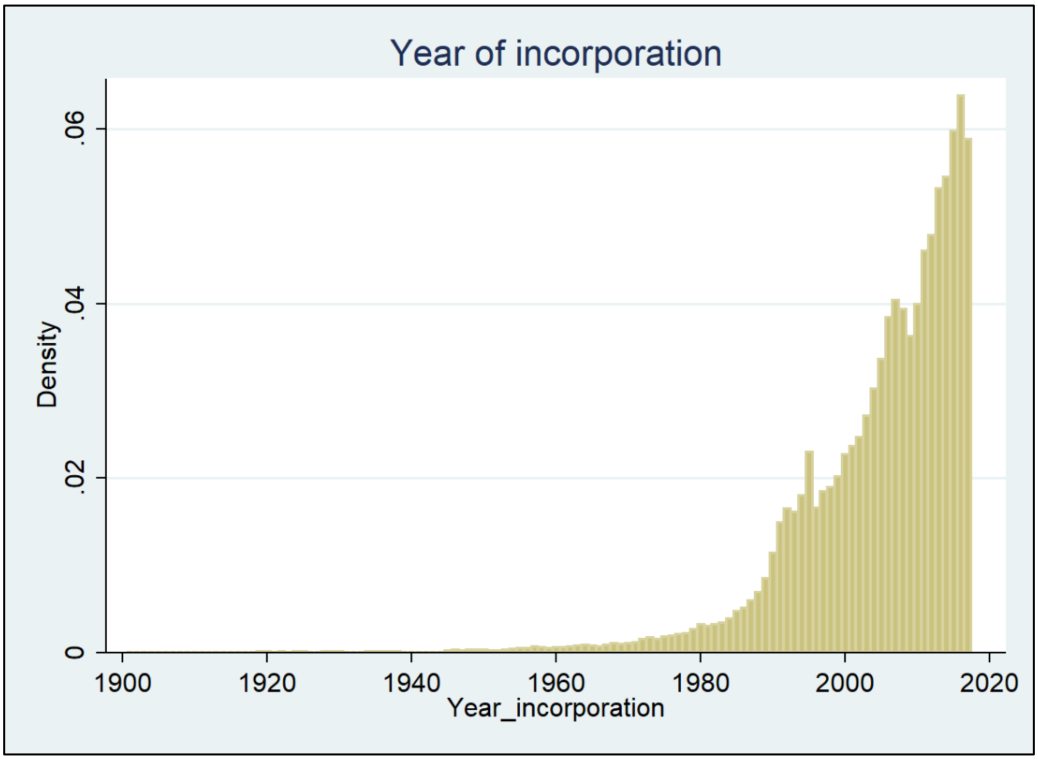

Figure 2.3 – Distribution of firm' years of incorporation over time

In fact, the graph shows an increasing trend of company foundations over time with their maximum around 2017. 50% of companies in the dataset were founded after 2009 and more the 90% of them after 1992. This distribution naturally shifts the availability of financial data towards the most recent years of the period considered.

Even if an actual multisectoral approach is not adopted, industrial sectors are a significant information exploited in the model definition process in next section. Thus, it may be useful to focus on distribution of firms among industrial sectors. The adopted criterion is the NACE Rev. 2 classification by Eurostat. It is the statistical classification of economic activities in Europe and its use is mandatory within Member States[16]. This classification is stratified in more and more specific layers. The lowest level of detail in grouping industrial activities is the one used in this thesis and reported in the following table.

NACE Rev. 2 sections

Section	Description
A	Agriculture, forestry and fishing
B	Mining and quarrying
C	Manufacturing
D	Electricity, gas, steam and air conditioning supply
E	Water supply, sewerage, waste management and remediation activities
F	Construction
G	Wholesale and retail trade, repair of motor vehicles and motorcycles
H	Transportation and storage
I	Accomodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M	Professional, scientific and technical activities
N	Administrative and support service activities
O	Public administration and defence, compulsory social security
P	Education
Q	Human health and social work activities
R	Arts, entertainment and recreation
S	Other service activities
T	Activities of households as employers, undifferentiated goods- and services-producing activities of households for own use
U	Activities of extraterritorial organisations and bodies

Table 2.2 – NACE Rev.2 industrial sector classification (upper level).

The next figure shows the distribution of firms among industrial sectors in the financial database. In order to simplify the understanding, sectors from M to U are grouped together as “Services” of various nature.

[16] https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

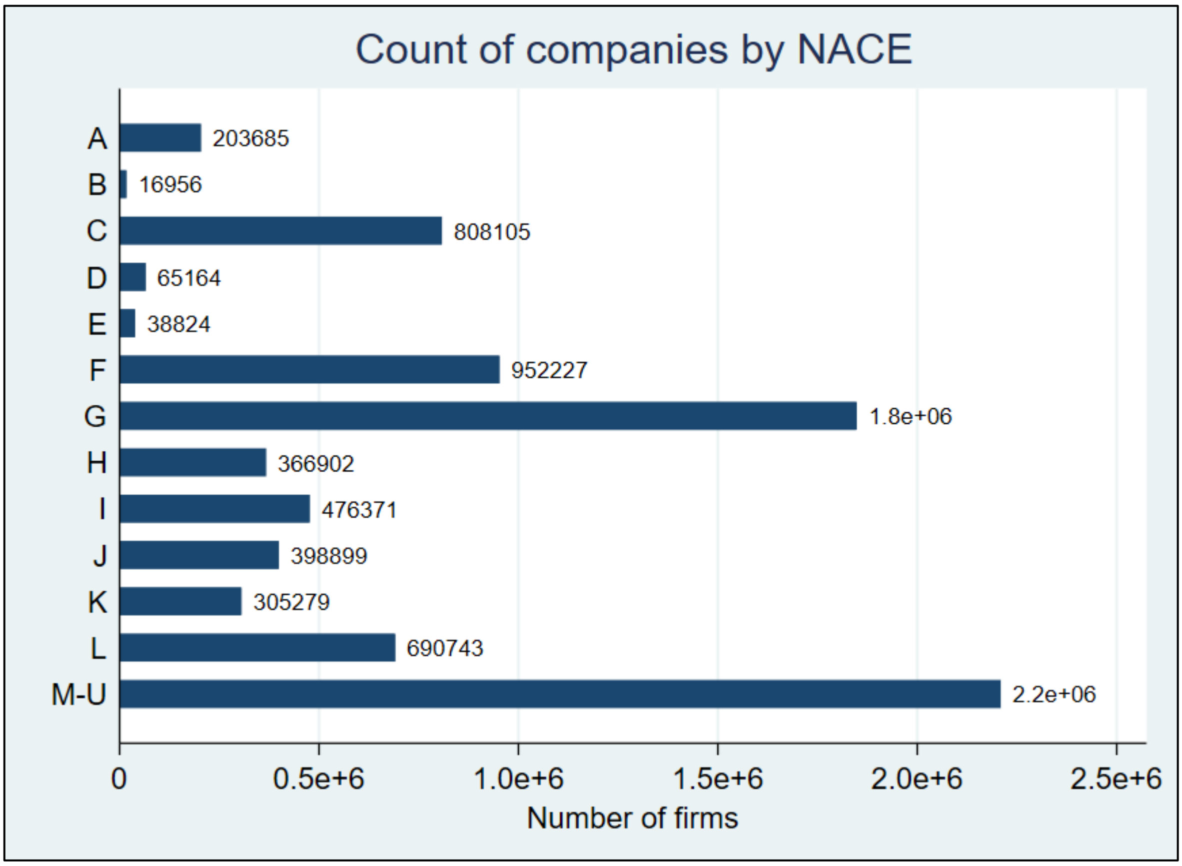

Figure 2.4 – Count of companies grouped by NACE Rev.2 industrial sector classification.

The graph clearly shows how the distribution of companies among sectors is heterogeneous. At the same time, each sector is well represented with the least populated one made of more than fifteen thousand firms. Thus, even if not part of a proper cross sectoral analysis, this classification can be meaningfully used for supplementary operations (such as a criterion for control group matching).

The next graphs represent distributions of variables to be studied in the panel dataset (long shaped). Lines show the shape of normal distribution and kernel density estimation, a linear way to represent density distribution useful to directly compare it with Gaussian density. At a first look, it may be noticed how the first three graphs, referring to logarithmic variables, all resemble closely normal distribution. While, on the contrary, ROA presents high-tail distribution compared to other variables mostly due to the fact that it is not logarithmic, even if winsorized.

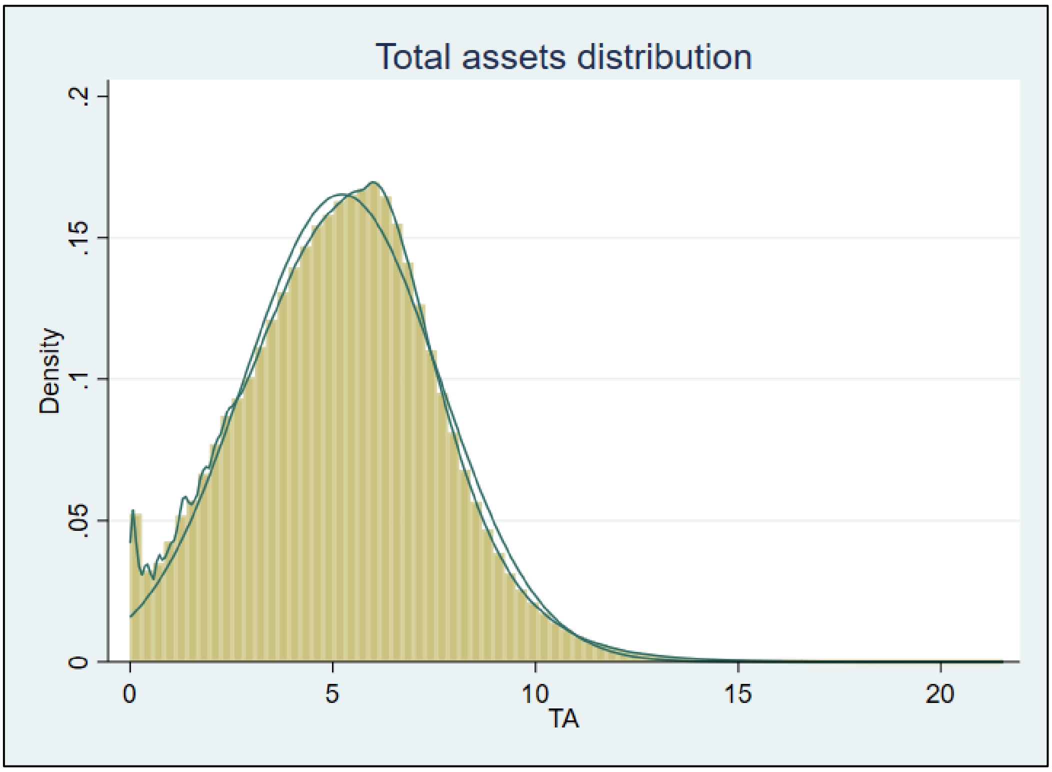

Figure 2.5 – Distribution of TA in the panel dataset.

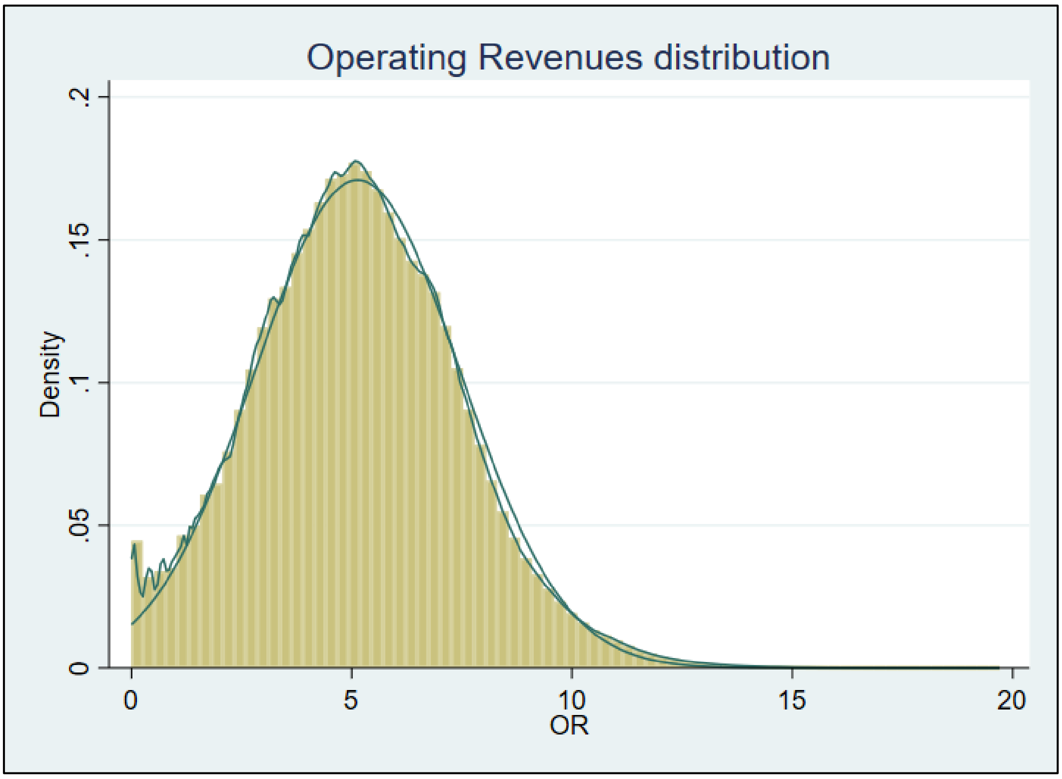

Figure 2.6 – Distribution of OR in the panel dataset.

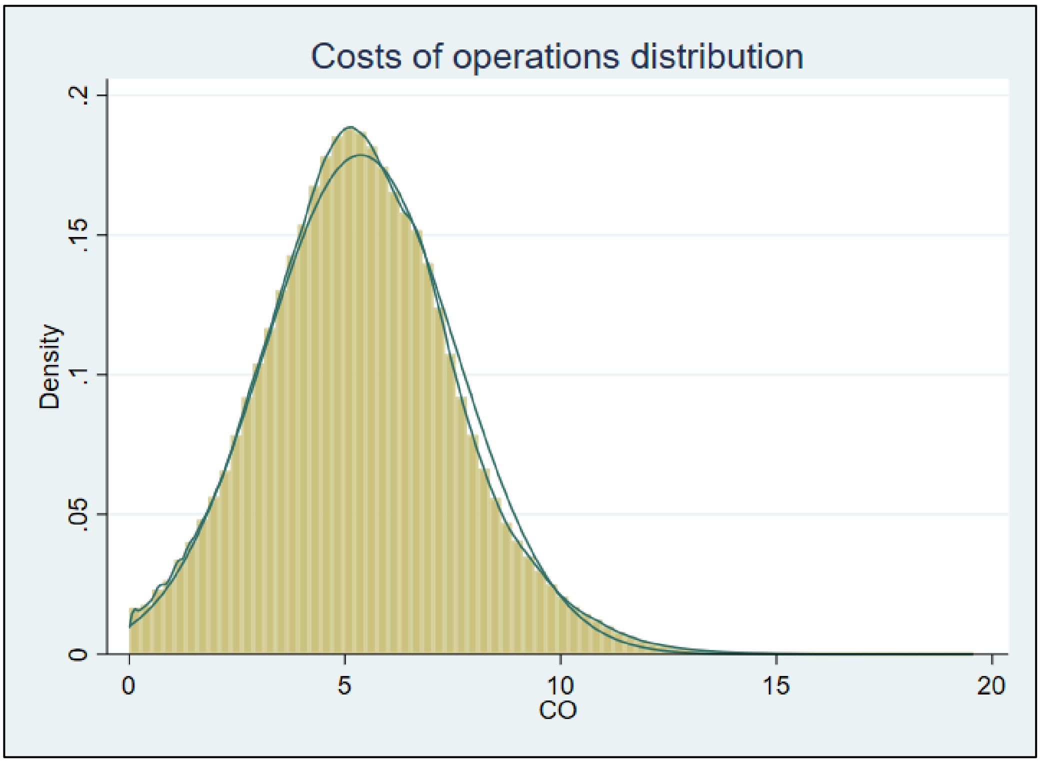

Figure 2.7 – Distribution of CO in the panel dataset.

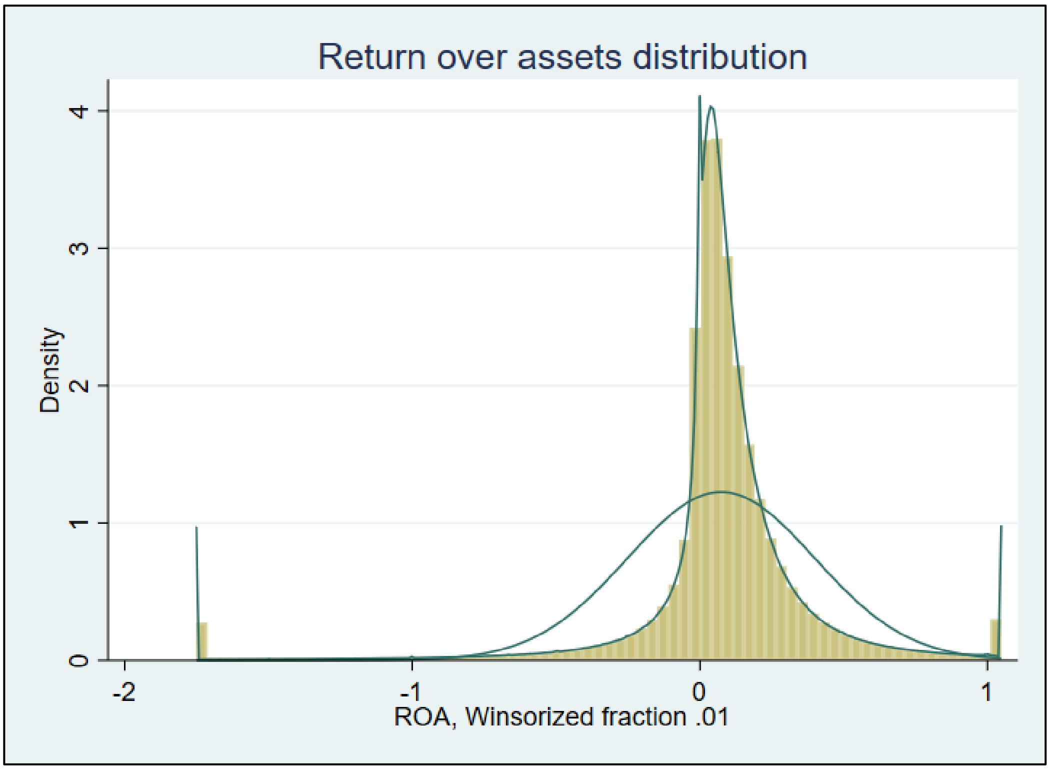

Figure 2.8 – Distribution of ROA in the panel dataset.

The next table reports summary statistics of the financial variables to integrate distribution comprehension. TA, OR and CO present very similar distributions, as reflected in percentiles, while ROA has a high standard deviation, more than four times the mean value, confirming its high-tail distribution behaviour.

Financial variable summary

	TA	OR	CO	ROA
Percentiles				
5th	1.238374	1.306439	1.756132	-0.32204
25th	3.559928	3.527236	3.854708	0.005061
50th	5.259524	5.076872	5.290598	0.070175
75th	6.789015	6.660575	6.77079	0.172535
95th	9.162829	9.061492	9.19482	0.519075
Mean	5.225323	5.132752	5.362707	0.074311
St. Dev.	2.412678	2.333647	2.233667	0.325645
Variance	5.821014	5.445907	4.989269	0.106045
Skewness	0.226835	0.269529	0.302057	-2.0364
Kurtosis	3.249961	3.181963	3.308978	15.20333

Table 2.3 – Table reporting summary statistics for financial variables.

Finally, two further statistical measures are considered. Skewness is an indicator that represents the symmetry of a distribution (with perfect symmetry reported for 0 skewness). It returns a light right asymmetry for logarithmic variables, while ROA presents a slightly left propended distribution (with left tails longer than right tails). On the other hand, kurtosis index measures how far the distribution is from normal one. As observed from the direct comparison between normal and kernel lines directly on graphs, this indicator confirms how logarithmic variable distributions are very close to normal (kernel=3), while ROA long tails keep it far from Gaussian distribution.

2.5.2 Climate data description

The EWE dataset presents several descriptive approaches. It is going to be described in its complete form even if some years of the period and some categories of events will not be considered in the final model.

Starting from the time distribution during the period considered, regardless of the country they affect, a slight increase in the number of events per year can be noticed.

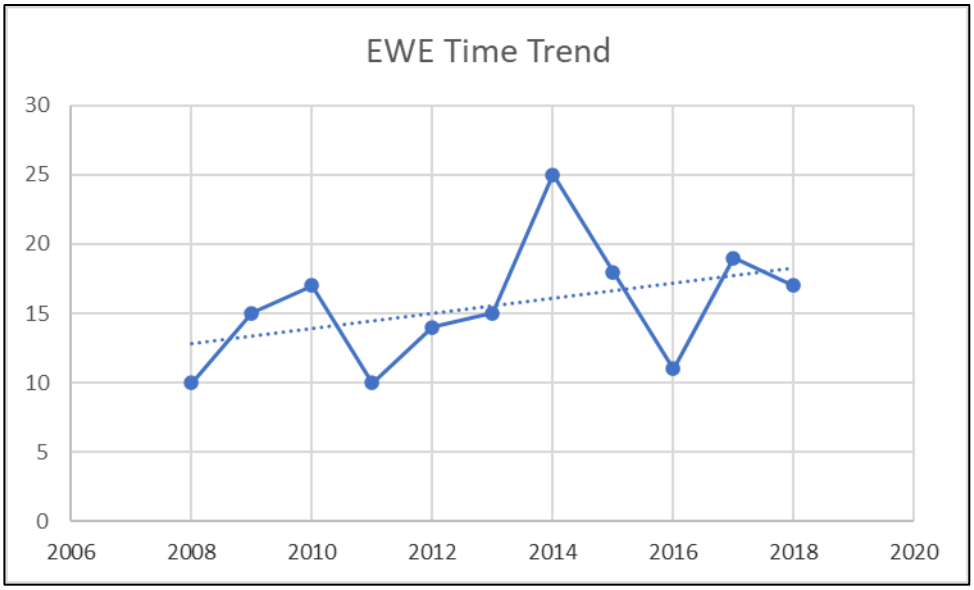

Figure 2.9 – EWE distribution over the time period.

With a minimum in 2008 and 2011 and a maximum for 2014, the overall mean is 15,5 events per year. The total count of unique events across the whole period is 171 events.

Among these events there are five different categories distributed as follows in the pie chart.

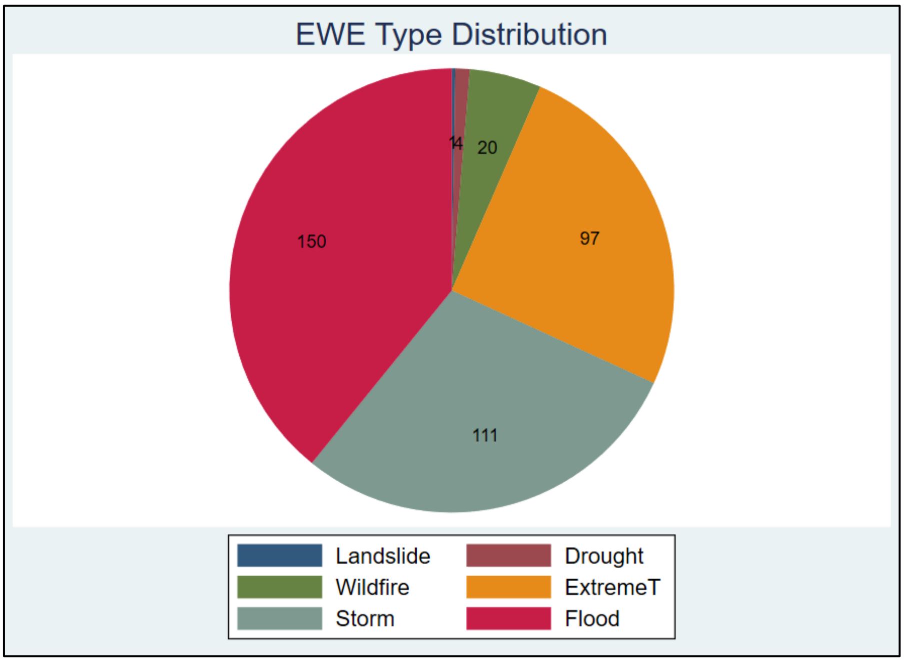

Figure 2.10 – EWE distribution by event category.

It can be noticed how Storm, Flood and Extreme Temperature events outnumber other categories.

Moreover, it may be interesting to observe their distribution among countries. As mentioned before, the total amount of unique events is 171, but the same event can affect more than one country. When it happens, the event is reported in all countries that affects. For this reason, it seems that the number of events in the next table increases with respect to the overall, but it is only the consequence of the frequent cross country nature of these events.

EWE country distribution

Country	Flood	Storm	Wildfire	Extr. Temp.	Drought	Landslide	TOT
Albania	8	-	-	2	-	-	10
Austria	4	4	-	2	-	-	10
Belgium	3	9	-	4	-	-	16
Bulgaria	8	1	-	4	-		13
Croatia	7	1	1	3	-	-	12
Czech Republic	5	3	-	4	-	-	12
Denmark	-	3	-	-	-	-	3
Estonia	-	-	-	2	-	-	2
France	11	18	2	8	-	-	39
Germany	6	13	-	6	-	-	25
Greece	7	-	3	2	-	-	12
Hungary	4	2	-	3	-	-	9
Ireland	3	3	-	-	-	-	6
Italy	19	7	1	6	2	1	36
Latvia	-	-	1	1	-	-	2
Lithuania	1	-	-	4	1	-	6
Luxembourg	-	1	-	-	-	-	1
Macedonia	4	-	-	3	-	-	7
Montenegro	3	-	1	1	-	-	5
Netherlands	-	7	-	2	-	-	9
Norway	-	2	-	-	-	-	2
Poland	4	6	-	12	1	-	23
Portugal	3	6	5	3	-	-	17
Romania	12	2	-	6	-	-	20
Serbia	13	-	-	8	-	-	21
Slovakia	3	1	-	2	-	-	6
Slovenia	2	-	-	1	-	-	3
Spain	8	5	5	1	-	-	19
Sweden	-	1	1	-	-	-	2
Switzerland	-	7	-	3	-	-	10
United Kingdom	12	9	-	4	-	-	25
TOT	**150**	**111**	**20**	**97**	**4**	**1**	**383**

Table 2.4 – Table reporting EWE distribution by country and by event category.

It is interesting to notice how the above distribution is the result of the sum of several event categories. In fact, the event type distribution within a country is not usually balanced: countries tend to report events of few categories. Italy is the only country that presents all five categories of event.

Drought and Landslide events report very few observations. In fact, Landslide presents a single observation in the whole period, while Drought events are reported just by three countries of the sample. Furthermore, Droughts tends to be reported as affecting the whole country making them not useful for the model approach of this thesis that will be discussed in further sections.

In these regards, a geographical distribution of events among NUTS3 levels may effectively explain this concept.

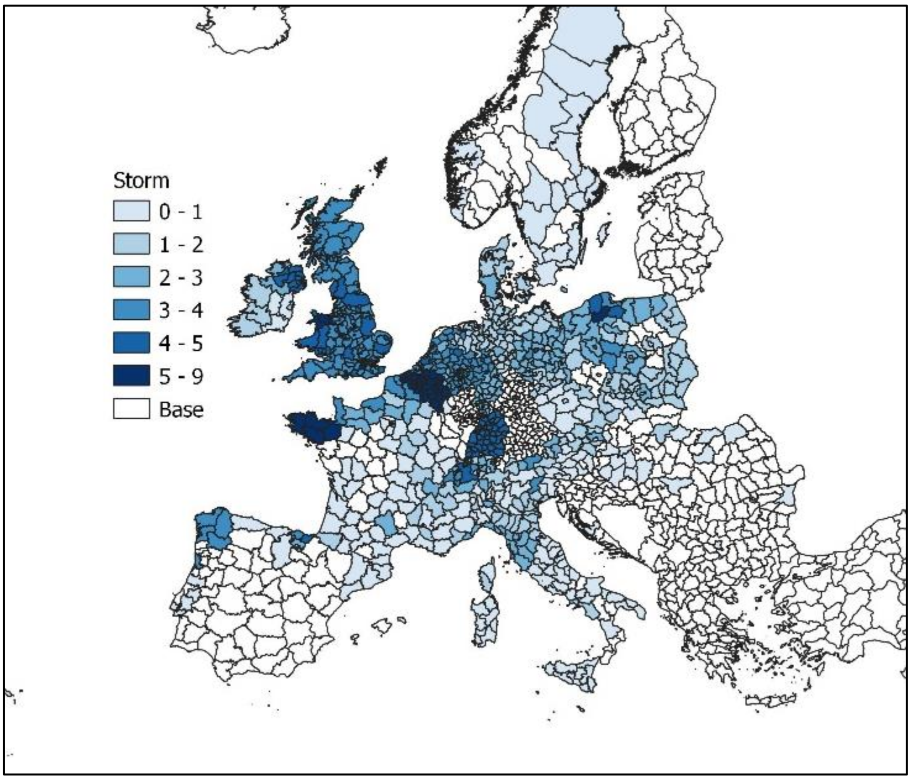

Figure 2.11 – Map of Storm distribution by NUTS3.

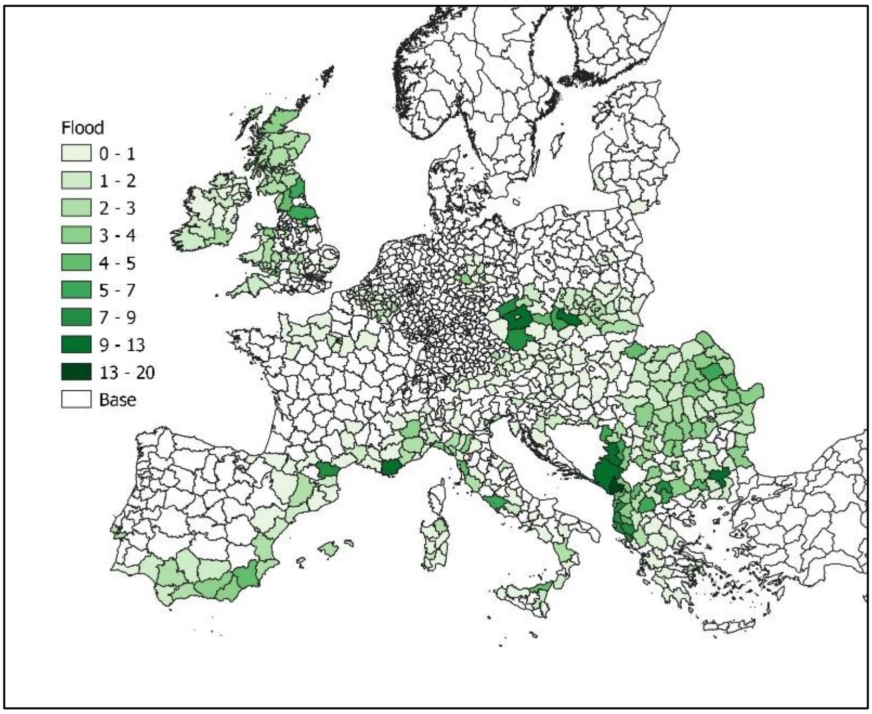

Figure 2.12 – Map of Flood distribution by NUTS3.

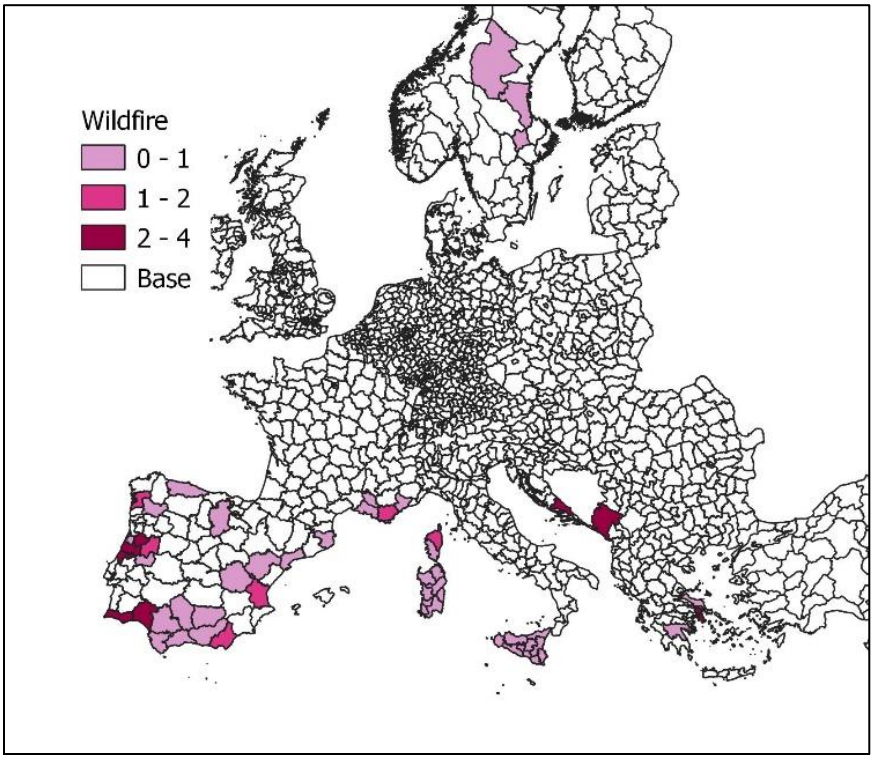

Figure 2.13 – Map of Wildfire distribution by NUTS3.

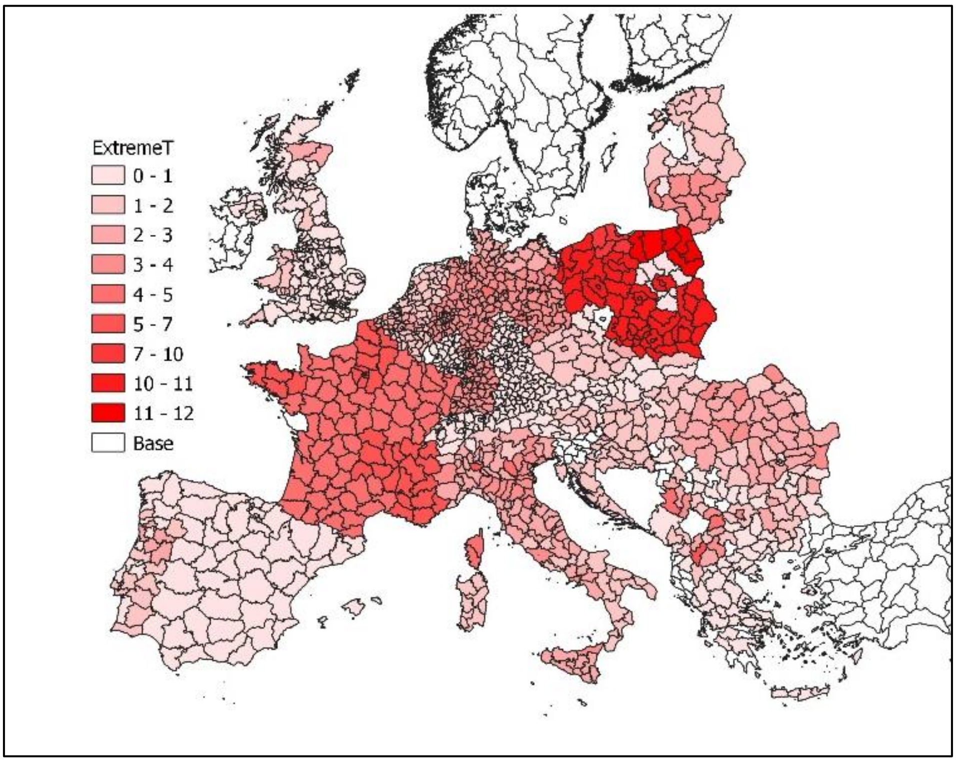

Figure 2.14 – Map of Extreme Temperature distribution by NUTS3.

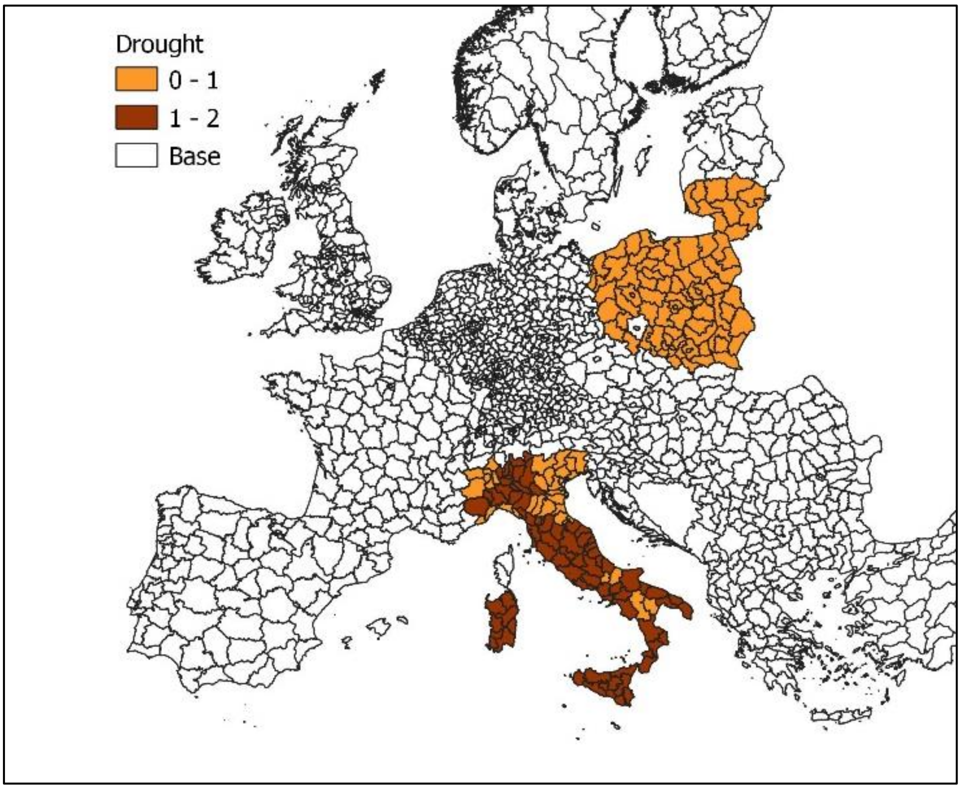

Figure 2.15 – Map of Drought distribution by NUTS3.

From these graphs, it can be noticed how Extreme Temperatures and Droughts generally affect very large portions of a country and often their whole surface. This peculiarity makes very difficult a within-country approach because they frequently present too many affected subjects with respect to unaffected ones, when not missing at all. On the contrary, other events present a suitable distribution for studying how these events can influence firm performance with respect to unaffected firms. Moreover, it is interesting to observe how differently these events are distributed in the map depending on their meteorological nature (e.g. wildfires in Mediterranean countries, storms in the north-west portion of the map).

Finally, it may be useful to observe the number of events that were lost in the matching phases described in previous sections. In fact, if all locations in the original dataset from EMDAT ended up unmatched, information on the event would be lost completely. Fortunately, as it may be seen in the following table, this eventuality manifested very few times, compromising negligible percentages in most of the countries with a significant presence of events.

EWE completeness

Country	Reported	Screened	Lost	Loss%
Albania	10	10	0	0%
Austria	10	10	0	0%
Belgium	16	16	0	0%
Bulgaria	13	13	0	0%
Croatia	12	12	0	0%
Czech Republic	12	12	0	0%
Denmark	3	3	0	0%
Estonia	2	2	0	0%
France	39	39	0	0%
Germany	25	23	2	8%
Greece	12	11	1	8%
Hungary	9	8	1	11%
Ireland	6	6	0	0%
Italy	36	36	0	0%
Latvia	2	1	1	50%
Lithuania	6	6	0	0%
Luxembourg	1	1	0	0%
Macedonia	7	7	0	0%
Montenegro	5	5	0	0%
Netherlands	9	9	0	0%
Norway	2	2	0	0%
Poland	23	23	0	0%
Portugal	17	17	0	0%
Romania	20	20	0	0%
Serbia	21	18	3	14%
Slovakia	6	6	0	0%
Slovenia	3	2	1	33%
Spain	19	19	0	0%
Sweden	2	2	0	0%
Switzerland	10	10	0	0%
United Kingdom	25	24	1	4%
TOT	**383**	**373**	**10**	**3%**

Table 2.5 -Table reporting lost EWE during the screening process compared to the original EMDAT database.

2.6 Data Analysis

This section is dedicated to the description of the analytical framework adopted to conduct the Difference-in-Difference (DID) estimation.

The DID model is the core of the framework, but several steps are needed before the actual DID can be performed. The analysis process can substantially be separated in two parts: the pre-estimation control group matching and the DID model. Next subsections are dedicated to the description of each of these parts.

2.6.1 Control group matching

The control group definition is an essential step for the DID estimation. When considering a treatment and the effects that it produces, two different groups can be defined: the treatment group, including all subjects that received the treatment, and the control group, consisting of all subjects that could have received the treatment but that did not. In this work, the treatment is the extreme event and the treated subjects are the affected firms, while the control group embraces all companies that were not affected but that present similarities with companies in the treated group.

It is important to define what similarity means and which variables are used to decide whether two companies are close enough each other. In the scope of this thesis, companies need to be compared under an economical and financial point of view. Thus, most of the variables involved to measure this similarity are those related to firm performance, but other characteristics must also be considered (such as country, sector, age).

To begin, the financial dataset is separated into several parts, each corresponding to a single country. This operation helped making the script lighter and it already selected the first control group condition: control group firms belong to the same country of treated firms. Then, each country file is merged to the first event file and the treatment dummy variable (A) is generated. A dummy variable is a variable that assumes the value 1 if a particular condition is satisfied, otherwise it equals 0. *A* equals 1 if the company is affected by the event and 0 if it is unaffected. After the dummy variable is assigned to each firm for a specific event, all firms incorporated after the event and all firms exceeding the Eurostat SMEs definition (Total assets>43M€ or Operating revenues>50M€) are dropped. The latter category was dropped because of the need of focusing on SMEs. At this point, two additional variables need to be generated: the average value of total assets (*mTA*) and the average value of operating revenues (*mOR*) calculated with observations of at least one year, till a maximum of three years, before the event (e.g. if event occurs in 2014, *mTA* is calculated as the average of total assets of 2013, 2012 and 2011; while if it occurs in 2011, *mTA* is calculated only with observations from 2010 and 2009). Finally, also *Year of incorporation* and *NACE Rev. 2* industrial sector classification are designated as discriminants for the control group.

All the control group discriminants are summarized in the following:

- *Country*: each country is processed separately, so both control and treated group firms belong to the same country.

- *Size*: not SMEs were dropped, so both control group firms and treatment group firms belong to the same size category.

- *Firm performance*: *mTA* and *mOR* represent the firm performance characteristics that control group firms need to match with treated firms.

- *Year of incorporation*: it represents the age of firms.

- *NACE Rev.2*: it is the industrial sector code classification.

All these characteristics contribute to the definition of the control group. The actual discrimination is computed through a propensity score matching process exploiting the STATA function *psmatch2*. The propensity score matching method (PSM) is a statistical procedure that estimates the probability of not treated subjects of receiving the treatment based on a list of independent variables. It is a useful technique to reduce the bias that could be generated by being treated. In fact, the outcome difference between groups may be dependent on factors that predict the treatment assignment rather than the treatment effect itself. The PSM reduces this bias by generating a sample of subjects that did not received the treatment, named control group, that is comparable to the treated group on all independent variables given as input. This method is particularly useful in treatment effect studies because it attempts to mimic randomization, while generally treatments are not randomly distributed.

In the specific case under study, extreme weather events are apparently randomly distributed treatments, so it may seem not very useful to adopt such method. However, even when accepting that meteorological events verify randomly, companies stricken by an extreme weather event share some characteristics related to the nature of the event. For example, a cyclone is more likely to affect coastal zones rather that inland regions, while flood risk is strongly affected by the presence of water. For these reasons, it is easy to imagine that a pattern, even if not always immediately clear, links companies that are likely to be affected by a particular category of event. The aim of the PSM procedure is to find unaffected companies that were likely to be affected by the event on the base of the independent variables given by input. In our specific case, PSM does not match companies on the base of events risk, but on the base of economic firm performance variables, with the goal of assigning to each treated firm at least one unaffected "twin" company. The group of "twin" companies matched this way is called the control group, while the batch of affected firms constitutes the treated group.

As before mentioned, the PSM was performed through the *psmatch2* STATA function. It is a function rich of features that implements several PSM methods such as one-to-one, k-nearest neighbors and radius matching. The function returns as output the propensity score, namely the conditional treatment probability, for each company, estimated on the base of independent variables given as inputs. For this thesis, the k-nearest neighbor method was chosen and it was designed 5 as a reasonable value for the number of neighbors. The 5-nearest neighbors method matches the treated firm considered to the 5 nearest unaffected companies ranking on the propensity score. This means that for each treated firm five untreated companies are selected.

The *psmatch2* independent variables (NACE Rev. 2, Year of incorporation, mTA, mOR) were set and, to avoid low scoring matches, it was also added the *caliper(0.9)* option to exclude propensity scores lower than 0.9.

Now that control and treated groups are defined, all not matched companies are dropped and the database is reshaped from wide to long in order to perform the DID model.

2.6.2 Difference-in-Difference model

Difference-in Difference is a statistical method adopted to measure the effect of a treatment on outcomes between a treated group and a control group over time. It compares the average change over time in the outcome experimented by the treated group with respect to the average change over time for the control group. Thus, it not only focuses on the treatment effect over time experimented by the affected subjects, treated group, but it also considers the outcome behaviour over time of an unaffected group of subjects, control group, and it measures the difference in outcome trends over time between the groups. The more accurate is the selection of the control group, the more unbiased and clear is the difference effect. For these reasons, while the treatment group is naturally selected, the control group matching needs particular attention to shape a batch of subjects as representative as possible, as described in the previous section.

The DID method can be performed through an Ordinary Least Squares (OLS) regression model.

$$y = \beta_0 + \beta_1 T + \beta_2 A + \beta_3 (T \cdot A) + \varepsilon$$

As in a classic linear regression, y is the main dependent variable, β_i are the estimation coefficients and ε is the stochastic error. T is the dummy variable for time, A is the dummy variable for groups and $(T \cdot A)$ is the combination of those dummy variables. The combined dummy coefficient (β_3) is what really measures the DID effect and it represents the main result of the regression for the scope of this thesis.

While *A* was already present in the dataset, *T* is generated on purpose. As before mentioned, the dataset, after the control group matching, was reshaped from wide to long. Thus, each row contains information about a company in a different year (10 rows for each company in the 10 years period under study). Starting from this new dataset form, *T* equals 1 for years that are greater or equal to the event year, while it equals 0 for years lower than the event year.

Dummy variables are used to select between different groups (affected/unaffected) and periods (before/after the event) and they are structured as follows:

- *T=1* : it identifies after treatment period.
- *T=0* : it identifies before treatment period.
- *A=1* : it identifies affected subjects.
- *A=0* : it identifies unaffected subjects.

Other variables are also generated as control variables, with the aim to reduce the bias that their omission could lead to. These control variables check for the effect of NACE Rev.2 industrial sector ($nace_j$), for year of the observation ($year_i$) and for the age of the firm (*age*). Specifically, while *age* variable contains the age of the firm generated by subtracting the year of the observation with the company's year of incorporation, $nace_j$ and $year_i$ are dummy variables. Thus, *i* assumes values from 2009 to 2018, while *j* goes from A to U (each letter representing a different industrial sector). These are the last variables needed to implement the final model equation.

$$y = \beta_0 + \beta_1 time + \beta_2 A + \beta_3 did + \beta_4 age + \sum_i \beta_{5i} year_i + \sum_j \beta_{6j} nace_j + \varepsilon$$

Where:

- y : the output variable under study *(TA, OR, CO or ROA)*.
- *time* : the dummy variable related to time *(T)*.
- A : the dummy variable representing treatment.
- *did* : the combined variable that identifies the DID effect $(T \cdot A)$.
- *age* : the control variable related to the age of companies.
- $year_i$: the i control variable out of ten related to the year of observation.
- $nace_j$: the j control variable out of twenty-three related to firm's industrial sector.

This regression was then run, through *reg* STATA function, for each country section (31) of the financial dataset and for each event of the EWE dataset.

Chapter 3 Results

This Chapter is intended to summarize and discuss main model results from the DID analysis. To the best of author's knowledge, this is the first micro-founded study on EWE shocks over SMEs' firm performance at European level.

Because of its novelty, this chapter has not to be intended as a definitive answer on this topic as well as it does not demand to be recognized as universally valid when referring to the interaction between extreme weather events and firm performance. Estimates obtained from the model analysis must be considered within their geographical, temporal and methodological context. For these reasons, the reader is advocated to caution when addressing this Chapter and he is strongly invited not to consider it out of the context of this thesis.

Finally, because of the elevated number of regressions performed, full estimates could not be reported, even in the Appendix. For this reason, estimation results are available integrally upon request to the author at the following contact: tiziano.milazzo@mail.polimi.it.

3.1 Estimation results

In next sections, estimation results are considered as a whole and summarized in graphs and tables in order to report an overview on estimates as comprehensive and accurate as possible. Results are grouped and arranged by event category, one for each section.

The following table reports the number of actually tested events by category and by country.

Regressed events

Country	Storm	Flood	Wildfire
AT	2	2	-
BE	3	2	-
BG	1	8	-
CH	6	-	-
CZ	1	3	-
DE	10	5	-
DK	2	-	-
EL	-	7	1
ES	4	8	4
FR	14	11	1
HR	1	7	1
HU	2	3	-
IE	2	2	-
IT	6	15	-
LT	-	1	-
MK	-	3	-
NL	4	-	-
NO	1	-	-
PL	4	3	-
PT	6	2	5
RO	1	9	-
RS	-	10	-
SE	1	-	1
SI	-	2	-
SK	1	3	-
UK	4	9	-
TOT	**76**	**115**	**13**

Table 3.1 . Table reporting regressed events by country and type (it is valid for all variables with the exception of ROA that misses two observations).

3.1.1 Flood results

In this subsection, several graphs report increasingly ordered DID coefficients estimations for each output variable. As in following subsections, only significant coefficients are considered and analyzed when reporting results.

The significance of a coefficient is determined by the interaction between its confidence interval and the zero value, determining its sign probability. In fact, only coefficients that have at least 90% of probability to be of one determined sign (positive/negative) are considered useful results. On the contrary, coefficients that present more than 10% probability of assuming an opposite sign with respect to the one reported by their mean value are considered not reliable.

The next graph shows estimated coefficients, increasingly ordered, with their confidence intervals. In red are highlighted estimates that do not respect the significance criterion (90% of confidence interval of unique sign), while in green are represented all significant coefficients. Graphically, only coefficients that have at least 90% of their confidence interval only in one side of the graph (delimited by the zero line) are considered significant.

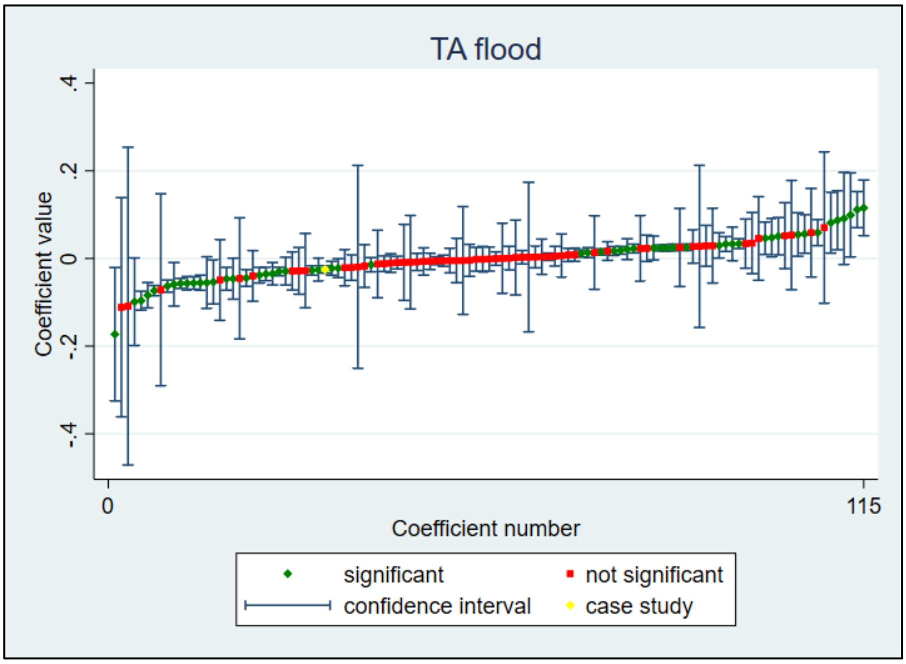

Figure 3.1 – Graph reporting flood DID coefficients for TA, with their confidence intervals.

To facilitate the understanding, from now on graphs of this kind will be reported without confidence intervals and with not-significant coefficients mean values reported as zeros. Furthermore, each graph also presents a coefficient, highlighted in yellow, representing the case study observation to be discussed in depth following in the section.

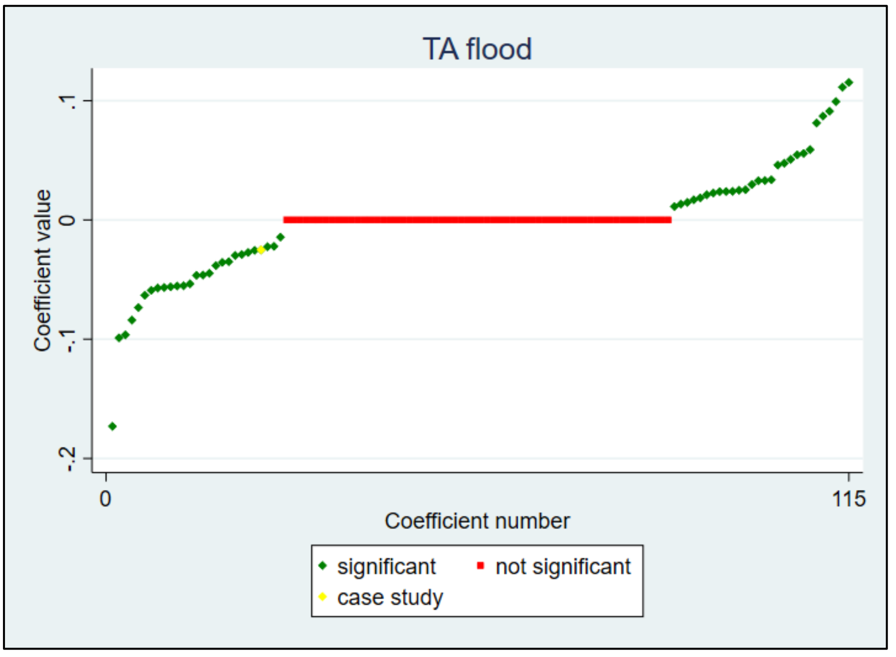

Figure 3.2 – Graph reporting flood DID coefficients for TA.

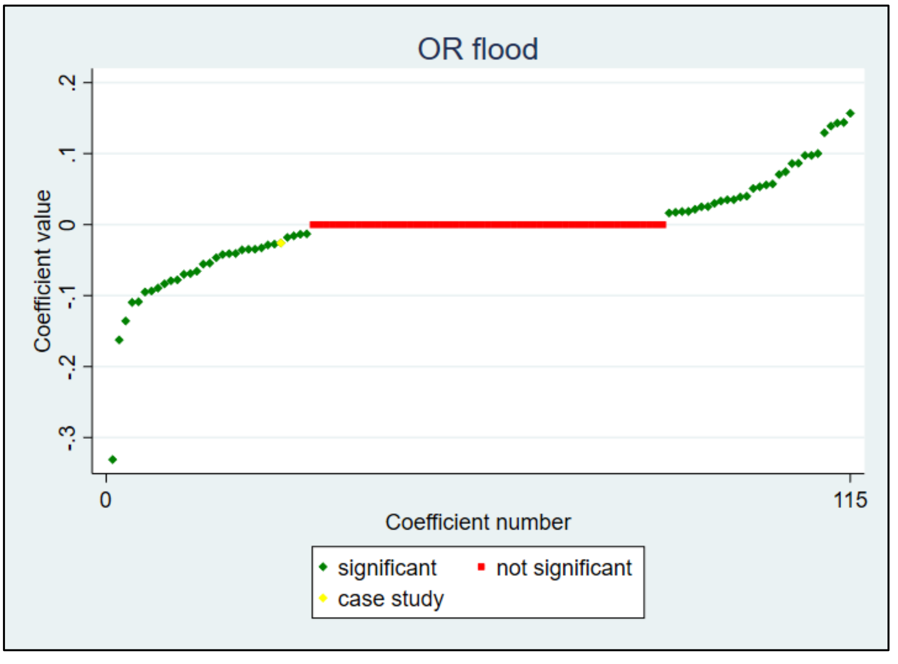

Figure 3.3 – Graph reporting flood DID coefficients for OR.

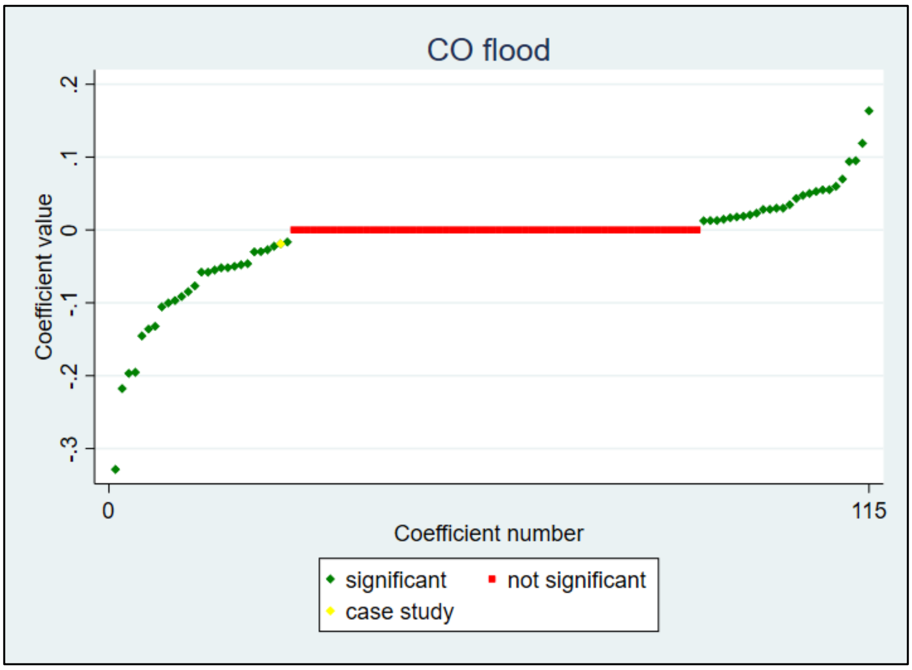

Figure 3.4 – Graph reporting flood DID coefficients for CO.

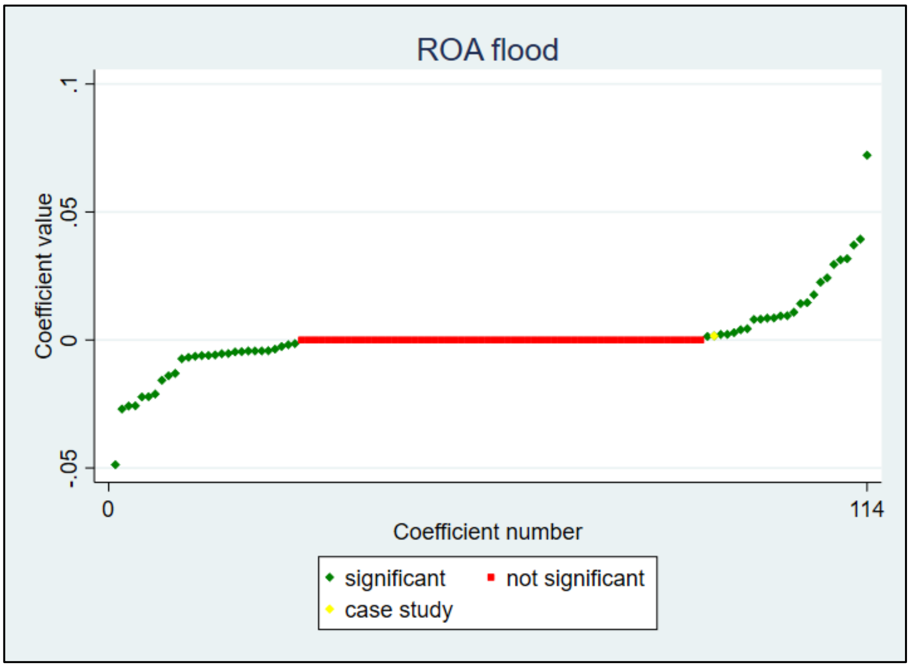

Figure 3.5 – Graph reporting flood DID coefficients for ROA.

It is immediately notable from graphs how the EWE effects seem not to be univocally directed on negative trends. At a first look, significant coefficients seem to equally distribute both on negative and positive signs. Next subsections will try to better explore results.

3.1.1.1 Coefficients summary

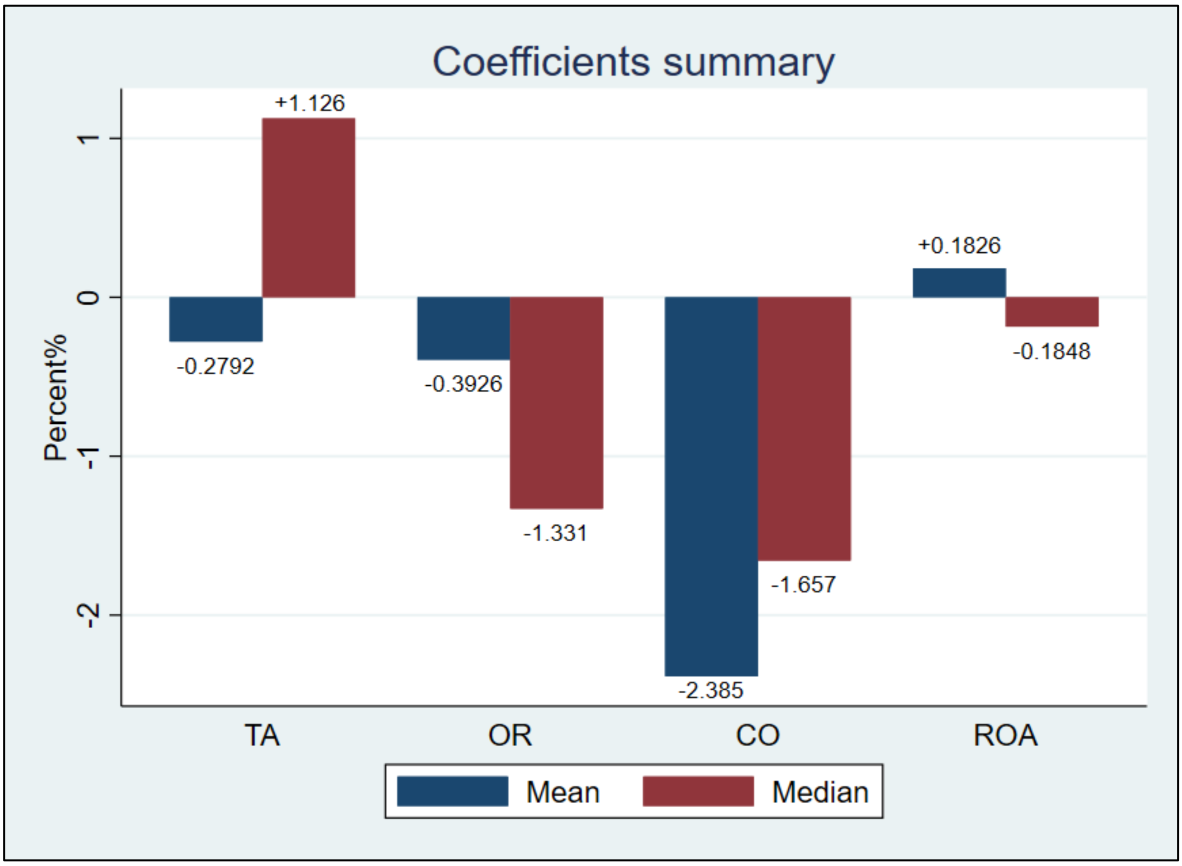

Figure 3.6 – Mean and median of flood DID coefficients by variable.

The previous graph reports mean and median of DID coefficients by variable. Only significant coefficients are considered to compute these values. Starting from this first graph some considerations can be raised:

- TA presents a slightly negative mean but a remarkable positive effect for the median, suggesting a negative skeweness of distribution.
- OR and CO show both concordant effects for mean and median, with CO reporting the highest absolute value effect (-2.4%).
- ROA produces the smallest effects and, such as TA, it has discordant trends between mean and median.

The next table reports the significant coefficients frequencies with respect to the overall number of coefficients. It clearly shows how the partition between significant and not-significant coefficients is mostly balanced (almost 50% for all variables).

Significance summary

Variable	Significant	Not-significant	TOT
TA	55	60	115
OR	60	55	115
CO	53	62	115
ROA	53	61	114

Table 3.2 – Count of significant and not-significant flood DID coefficients.

Finally, it is useful to observe significant coefficients density distribution. Through these graphs it is possible to better explain some considerations already made for mean and median values.

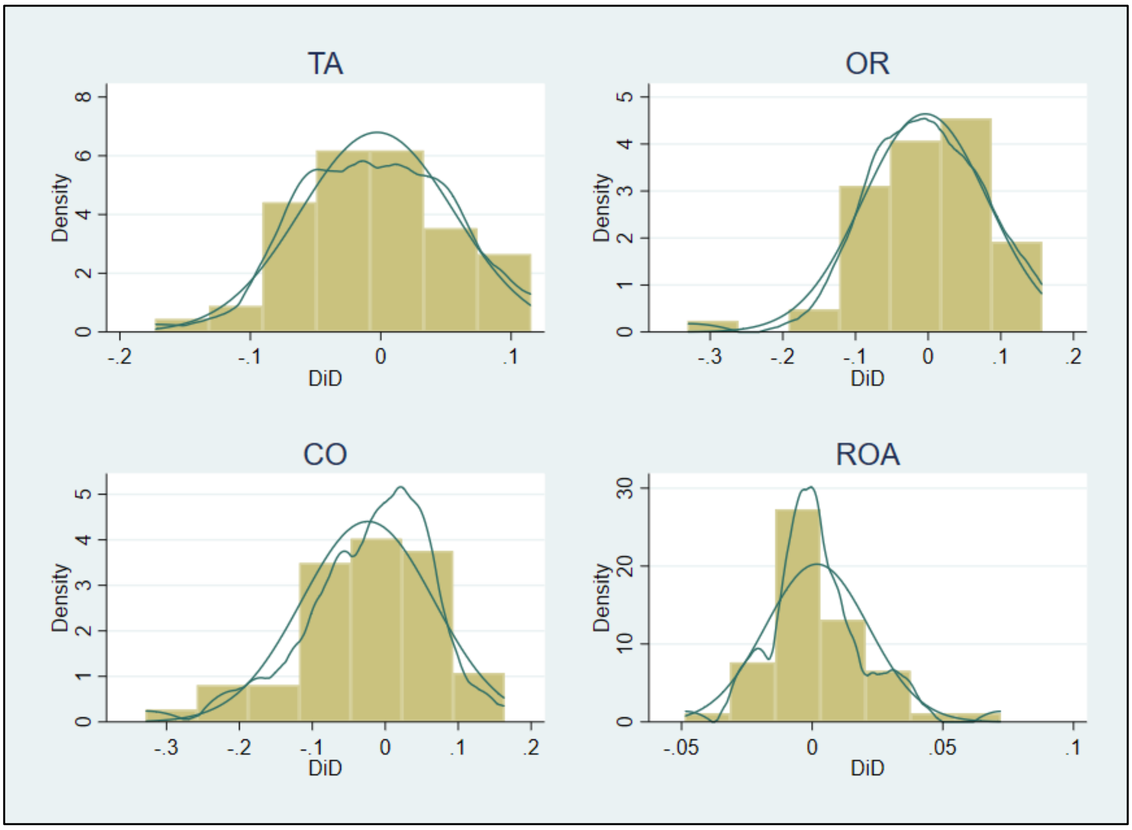

Figure 3.7 – Density distribution of significant coefficients values, by variable.

From these graphs, it is possible to make some final considerations:

- TA presents its highest density bars in the proximity of zero. Positive tails are more dense, but negative tails show more extreme values, leading the negative skeweness propensity.
- OR reports its highest density bar in the positive side. As well as TA, right tails are dense, while left tails assume extreme values.
- CO behaves similarly to previous variables, but the negative skeweness is more marked. In fact, most of the density is near the zero, but left tails are thicker and more extreme than right ones.
- ROA, differently from other variables, has thin tails with a positive skeweness and the density peak on the left side of the zero. This peculiar configuration moves the mean on a negative effect (because of the density peak), but the median on the positive effect (because of extreme positive tails).

3.1.1.2 Case study – Genova 9 Octorber 2014

During the night between the 9th and the 10th of October 2014 the Bisagno river overflowed its banks. In the following 24 hours the Genova province has faced a shattering storm that poured out 436mm of rain, more than a half times the monthly average (171mm). The day after, the Washington post reports: *"Cars piled up in the streets as a torrent of flood water swept them downstream. Video of the event shows water two to three feet deep rushing through Genoa's streets. (...) Crews clearing debris from the streets early on Friday found cars piled up on top of each other and sunk into huge holes in the roads. Thick layers of mud reached high up many shop walls."*[17] A man lost his life and damages amount to 300 milion euros, the city is on its knees and people feel forsaken. Aftermath rescues are mostly based on volunteers, they are called "Mud Angels", that help flooded dealers with shovels to free streets from the mud. The tension is so high that in some districts police is forced to intervene in pacifically quelling demonstrants' frustration: a similar flood struck the city just three years before and they blame the government of not having done enough in securing the population.[18]

The flash flood of Bisagno river was so dreadful and unlikely that researchers dedicated a paper on studing the fenomenon. In their work, Silvestro, Rebora et al. describe the event as an "an almost perfect flood. (with a peak flow) estimated to be a 100-200 years order return period". They investigated the effect of spatial and temporal pattern of rainfall on the river's streamflow demonstrating how difficult was to predict a flash flood of that nature with large anticipation (Silvestro, Rebora et al., 2016).

Anyway, unfortunate conditions and the unusual intensity of the fenomenon were not the only causes of the calamity. In fact, damages have been exacerbated by the unregulated building and mistreated public infrastructures summed together with the endogenous hydrogeologic risk of the area.

---

[17] https://www.washingtonpost.com/news/capital-weather-gang/wp/2014/10/10/torrential-rain-causes-deadly-flash-flooding-in-genoa-italy-video/

[18] Alluvione a Genova, è ancora allerta. Rabbia e barricate, gli 'angeli del fango' lavorano tra 300 milioni di danni - la Repubblica

In the following, it is reported the table containing DID estimation results of Genova flood. It is possible to notice how all DID coefficients are at their highest significance level (99% significance).

	TA	OR	CO	ROA
time	-0.1325259***	-0.2729764***	-0.0410157***	0.0044444***
	(0.0046845)	(0.0061402)	(0.0050550)	(0.0004325)
A	0.1511222***	0.2757683***	0.1802445***	0.0018218***
	(0.0030770)	(0.0042754)	(0.0036214)	(0.0002947)
did	-0.0250291***	-0.0260309***	-0.0188637***	0.0016224***
	(0.0043223)	(0.0060351)	(0.0050195)	(0.0004194)
age	0.0348354***	0.0259822***	0.0279595***	-0.0003699***
	(0.0000821)	(0.0001048)	(0.0000944)	(0.0000060)
N	3.06e+06	3.05e+06	2.59e+06	2.60e+06

*Table 3.3 – Regression estimators for relevant indipendente and control variables of Genova flood 2017. Standard errors are reported in brackets, N is the number of observations; * significant at p<0.1; ** significant at p<0.05; *** significant at p<0.01.*

All logarithmic variables present strongly negative coefficients, while ROA exhibits a mild positive trend. ROA may increase because it is the ratio of two financial variables that both, but differently, decrease EBITDA and Total assets.

3.1.2 Storm results

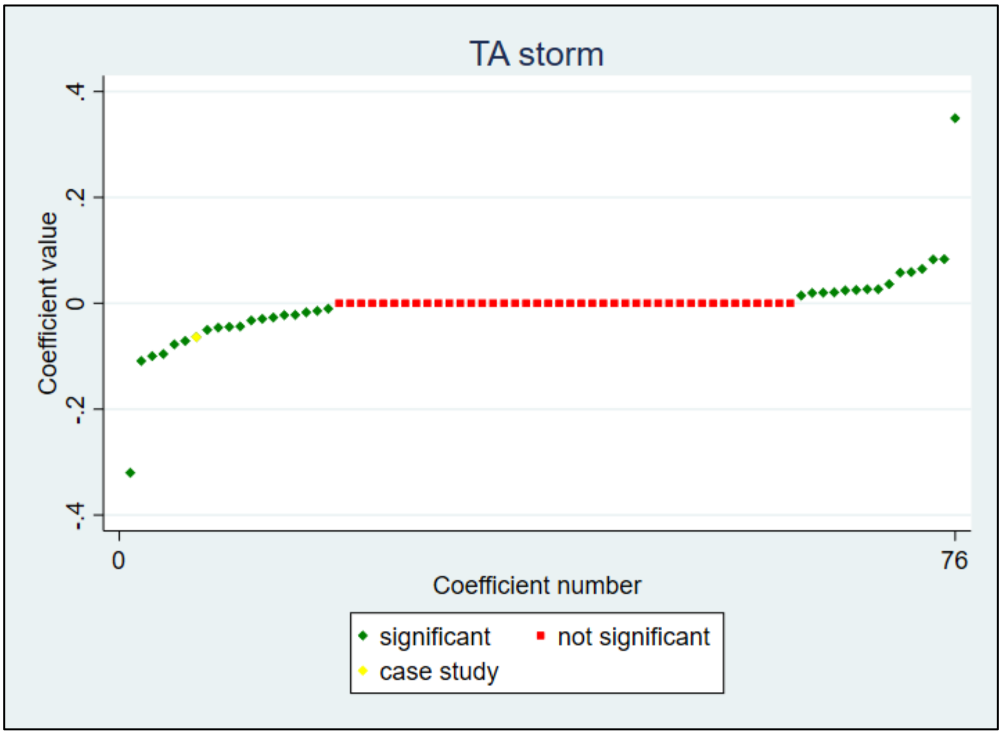

Figure 3.8 - Graph reporting storm DID coefficients for TA.

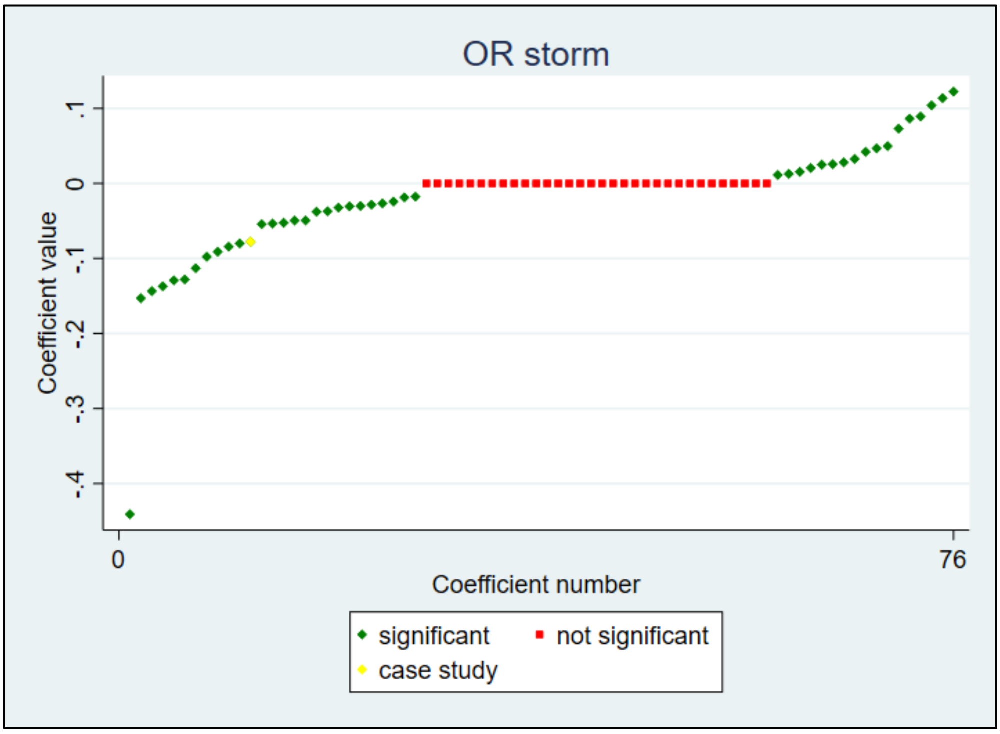

Figure 3.9 - Graph reporting storm DID coefficients for OR.

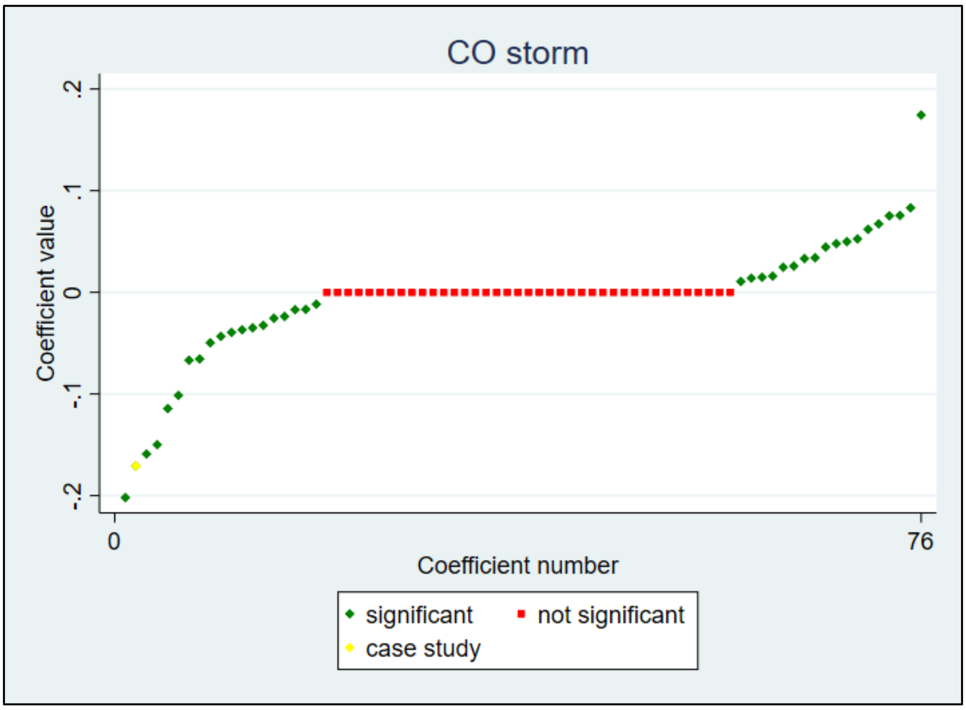

Figure 3.10 - Graph reporting storm DID coefficients for CO.

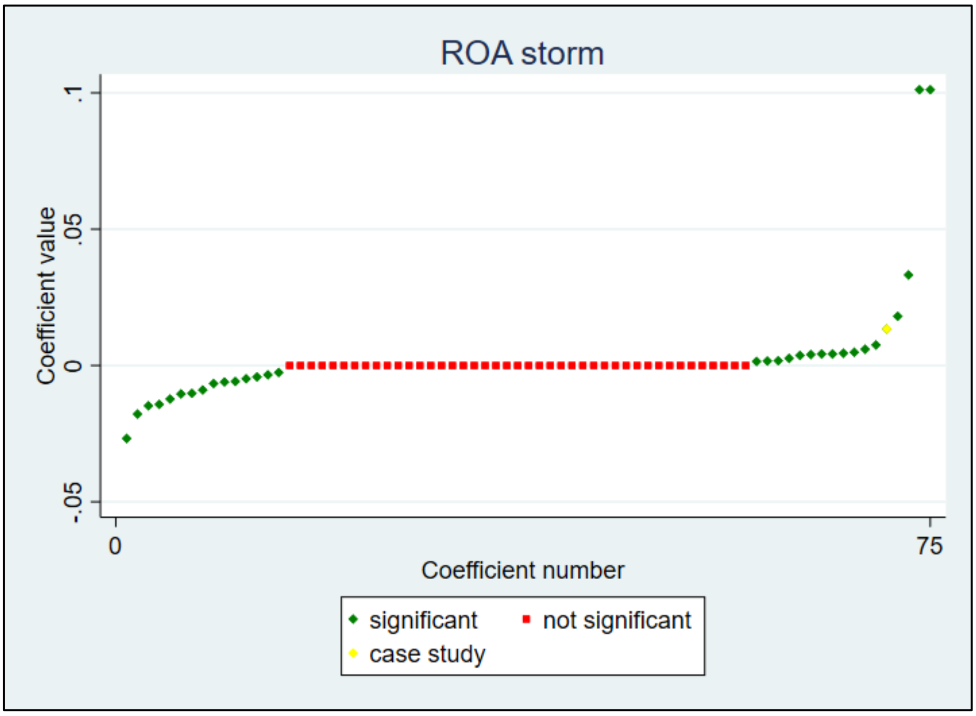

Figure 3.11 - Graph reporting storm DID coefficients for ROA.

Exactly as for Floods, Storm coefficients seem, at a first look, to equally distribute between positive and negative side.

3.1.2.1 Coefficients summary

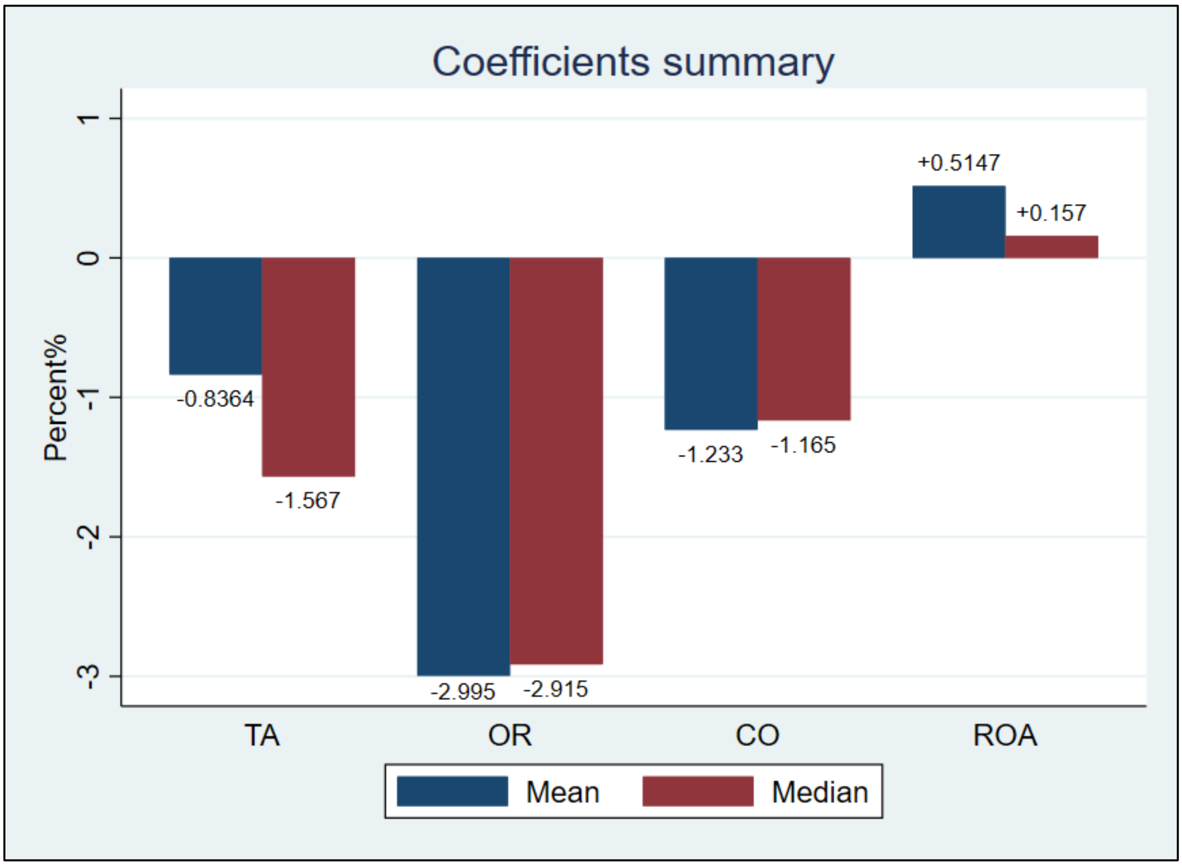

Figure 3.12 – Mean and median of storm DID coefficients by variable.

This graph reports mean and median of coefficients with the same criterion applied for Floods: only significant coefficients are reported. Starting from the graph, some considerations can be raised:

- TA exhibits both negative mean and median, the median value is higher suggesting thicker left tails.
- OR and CO behave similarly. With very similar values of mean and median that may indicate a symmetrical distribution. Additionally, OR presents values of coefficients more than double with respect to CO.
- ROA shows both positive mean and median, reflecting a mild increase of EBITDA weight over Total assets.

The next table reports the significant coefficients frequencies with respect to the overall number of coefficients. As well as Floods, the table shows a very balanced partition between significant and not-significant coefficients.

Significance summary

Variable	Significant	Not-significant	TOT
TA	34	42	76
OR	44	32	76
CO	37	39	76
ROA	32	43	75

Table 3.4 - Count of significant and not-significant storm DID coefficients.

Finally, it is reported significant coefficients density distribution.

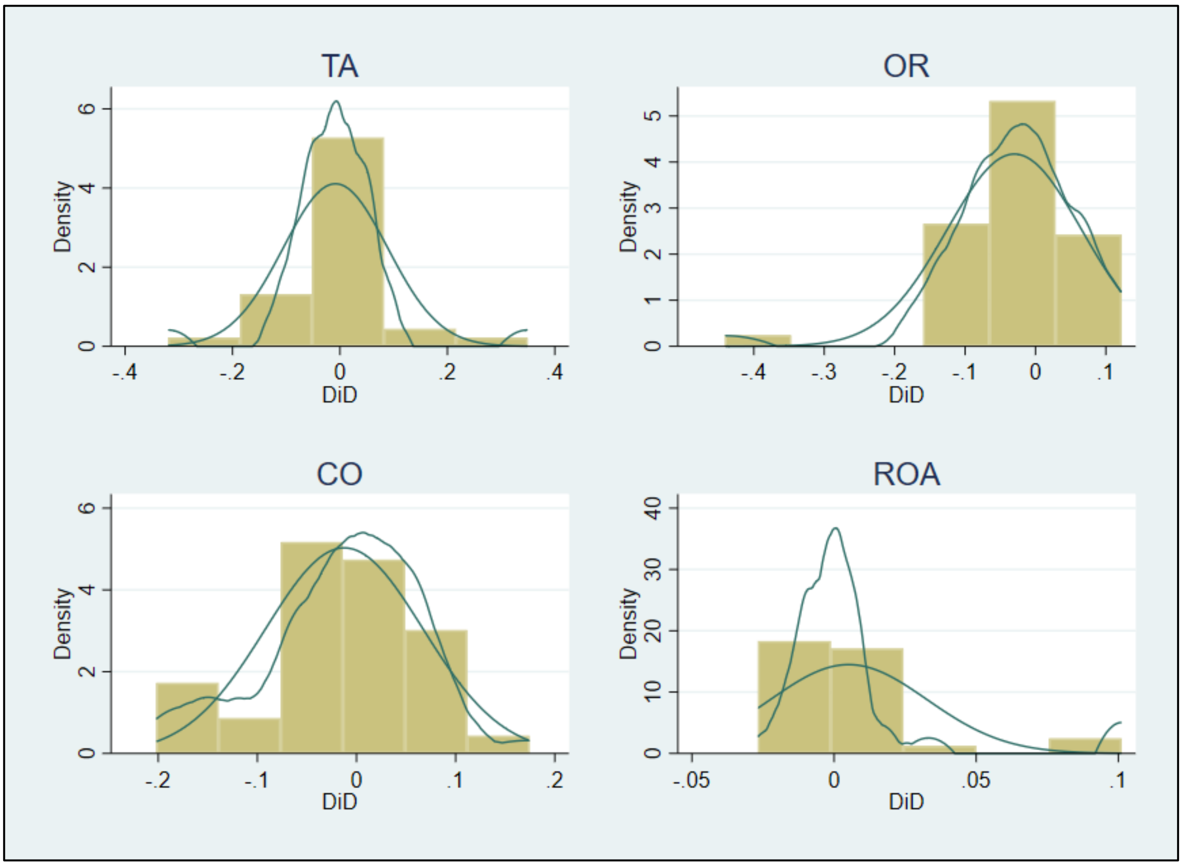

Figure 3.13 – Density distribution of significant coefficients values, by variable.

From this graph, the following considerations are raised:

- TA exhibits a very symmetrical distribution with similar thin tails. The mean value is shifted on the left by the slightly thicker density in the negative side close to zero.
- OR reports a very symmetrical distribution too, but the very extreme left tail severely affects the mean value.
- CO shows a peculiar distribution. In fact, there are high density bars in the positive side, but the peak on the close left of the zero and the thick left tail places the median effect firmly on the negative side.
- ROA presents a pretty flat distribution around zero with its mean value conditioned by the right tail extreme.

3.1.2.2 Case Study - North Rhine-Westphalia 9 June 2014

The 9th of June 2014 western Germany was hit by a dreadful storm. The most affected region was North Rhine-Westphalia where six people died because of falling trees and other incidents caused by the weather. The storm just followed the hottest weekend of the year with temperature peaks of 36°C and winds were recorded to blow up to 150 km/h at the Duesseldorf airport. Falling trees blocking streets and damages at power lines were reported, several railway lines were suspended and the Duesseldorf airport was even forced to interrupt the service for an hour.[19] BBC news reports the words of the North Rhine-Westphalia's Interior Minister Ralf Jaeger commenting the event: *"We must reckon that the total damage will run into double-digit millions" adding "That was one of the worst storms to hit (the region) in the past 20 years"*[20].

In the following, it is reported the table containing DID estimation results of North Rhine-Westphalia storm. All DID coefficients are significant, but at different levels.

	TA	OR	CO	ROA
time	0.3128972***	-0.3557071***	0.7268382***	-0.0076938
	(0.0159105)	(0.0464338)	(0.0712133)	(0.0090216)
A	-0.0187086	0.0893481***	0.1884018***	-0.0003278
	(0.0129671)	(0.0225136)	(0.0400664)	(0.0045961)
did	-0.0640403***	-0.0777212***	-0.1709127**	0.0134094*
	(0.0187950)	(0.0283282)	(0.0732456)	(0.0081477)
age	0.0110851***	0.0093840***	0.0135865***	-0.0000987**
	(0.0001644)	(0.0002192)	(0.0004864)	(0.0000387)
N	1.70e+05	8.46e+04	2.53e+04	2.80e+04

*Table 3.5 – Regression estimators for relevant indipendente and control variables of North Rhine-Westphalia storm 2014. Standard errors are reported in brackets, N is the number of observations; * significant at p<0.1; ** significant at p<0.05; *** significant at p<0.01.*

[19] La tempesta in Germania - Il Post
[20] Germany storms: Six dead in North Rhine-Westphalia - BBC News

As for Floods, logarithmic variables present strong negative coefficients, while ROA seems to reflect an increase of EBITDA weight over Total assets. It may be considered counter intuitive that the highest decrease is associated to a voice of cost (CO), but a cost decrease does not necessarily mean an increase in revenues, as reflected by the OR coefficient. In fact, costs may also decrease because of the disruption of assets and reduction of operations that force the firm to downsize their business.

3.1.3 Wildfire results

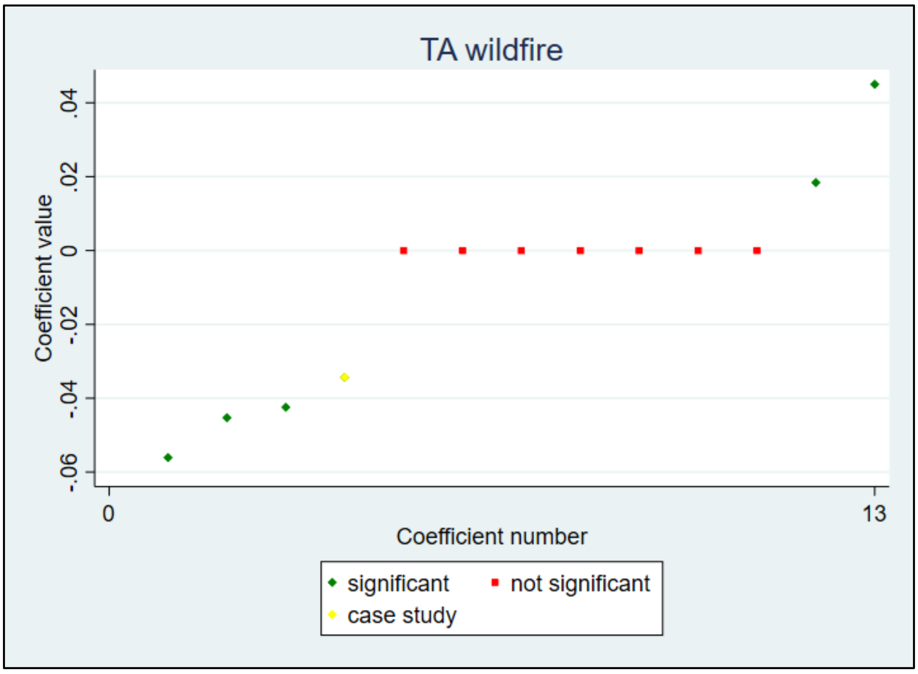

Figure 3.14 - Graph reporting wildfire DID coefficients for TA.

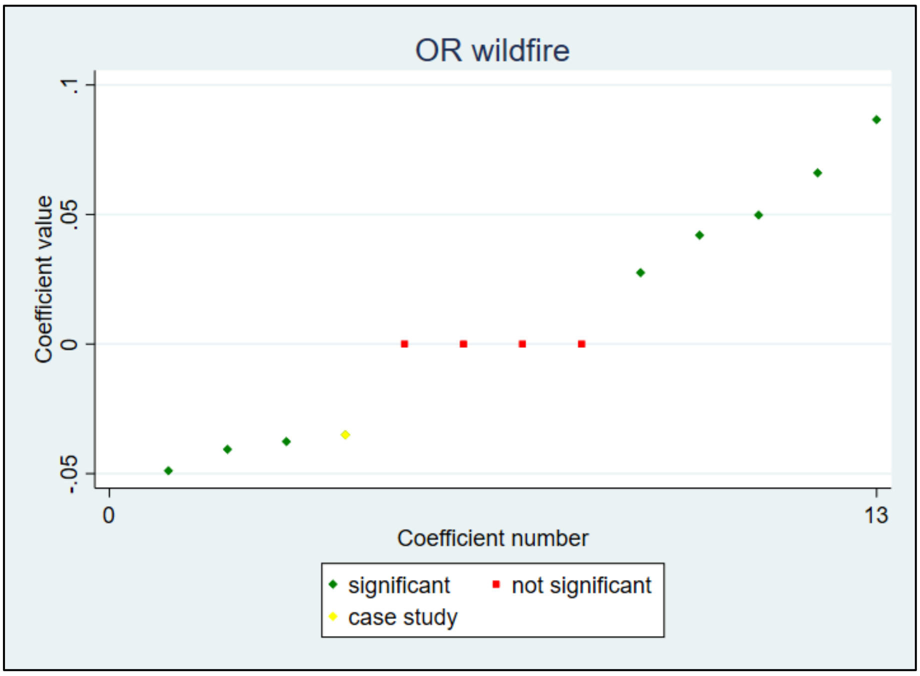

Figure 3.15 - Graph reporting wildfire DID coefficients for OR.

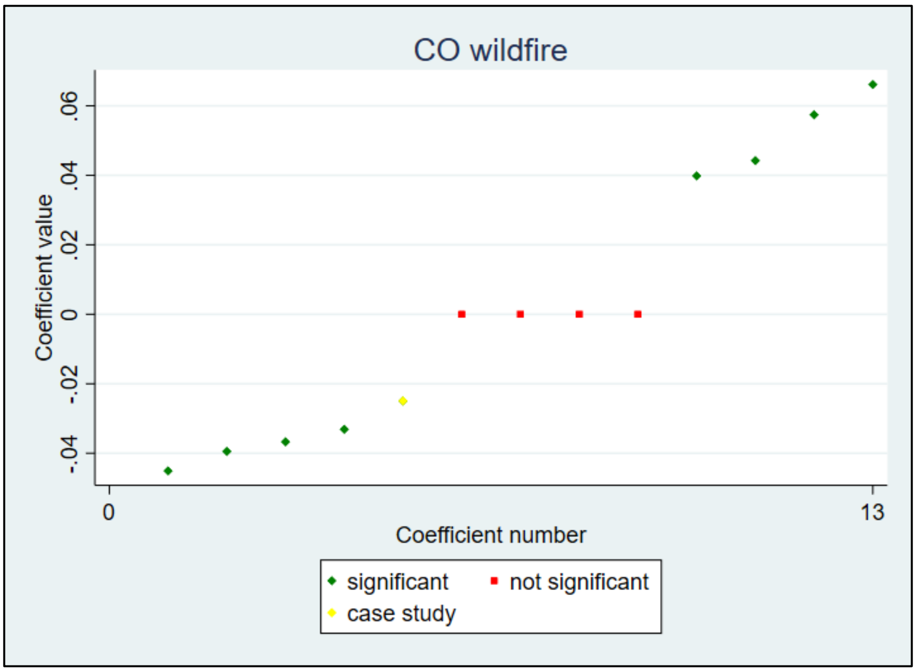

Figure 3.16 - Graph reporting wildfire DID coefficients for CO.

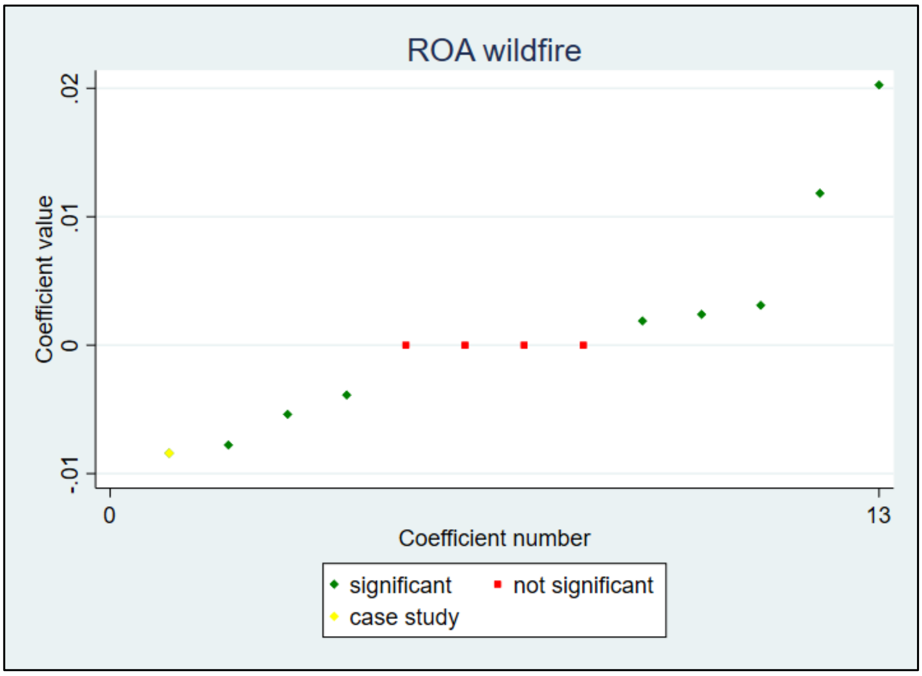

Figure 3.17 - Graph reporting wildfire DID coefficients for ROA.

From these graphs, it is immediately possible to notice how Wildfire observations are far less than other events considered. Nevertheless, also Wildfires present coefficients on both sides of the graph (positive/negative).

3.1.3.1 Coefficients summary

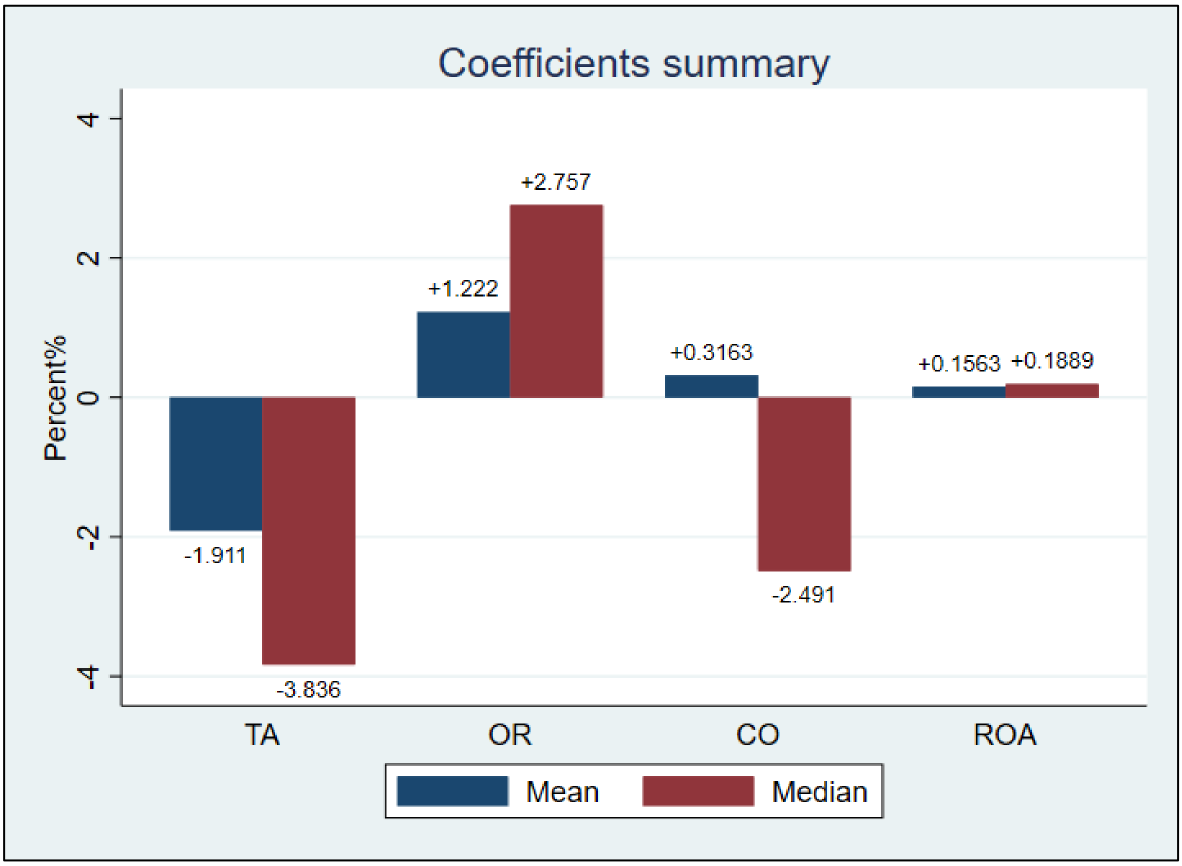

Figure 3.18 – Mean and median of wildfire DID coefficients by variable.

From mean and median graph of significant DID coefficients by variable some considerations can be raised:

- TA shows the highest mean and median absolute value effects respectively, both negative. The high median value suggests a strong left side density.
- OR exhibits positive values. Differently from TA, the strong positive median value reflects a right propensity of the density distribution.
- CO presents a mild positive mean but a highly negative median. This behaviour may originate from a dense negative distribution, but with extreme positive tails.
- ROA reports mild positive values for both indicators. This is probably due to symmetrical distribution with a slightly right side propension.

The next table reports the significant coefficients frequencies with respect to the overall number of coefficients. Differently from other events, the table shows a slightly propension towards significant coefficients.

Significance summary

Variable	Significant	Not-significant	TOT
TA	6	7	13
OR	9	4	13
CO	9	4	13
ROA	9	4	13

Table 3.6 - Count of significant and not-significant wildfire DID coefficients.

Finally, it is reported significant coefficients density distribution.

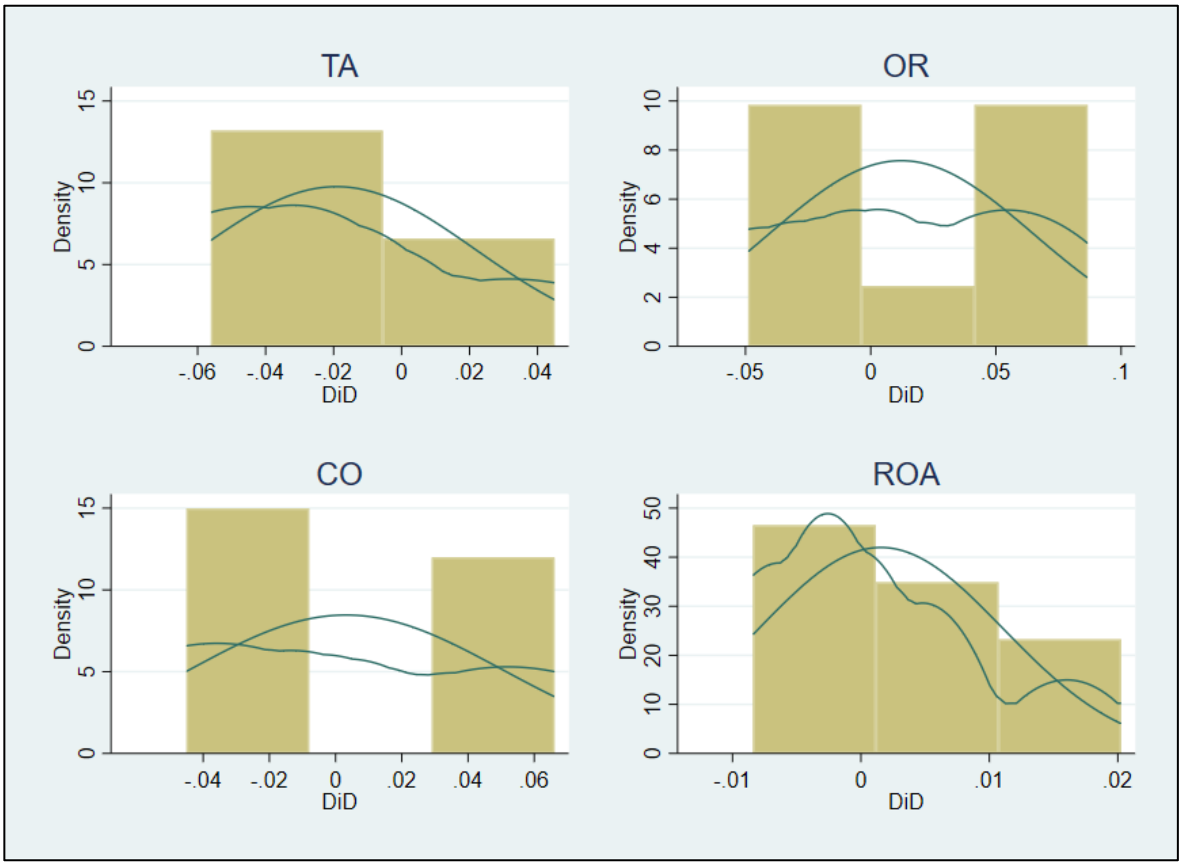

Figure 3.19 – Density distribution of significant coefficients values, by variable.

From this graph, the following considerations are raised:

- TA shows a symmetrical distribution around zero, but with an almost double high negative density.
- OR reports a symmetrical distribution too, with the lowest positive peak determining the overall positive mean trend.
- CO also presents a symmetrical distribution. In this case, the negative peak determines the negative median, but the more extreme positive bar shifts the overall mean value to the right side.
- ROA exhibits its density peak on the left side, but also a relevant thick tail on the positive one. The result is an overall slightly positive mean effect.

3.1.3.2 Case study Pedrógão Grande 17 June 2017

The 17th of June 2017 one of the worst wildfires of Portugal history raged through the central region of the country near Pedrógão Grande. Consequences were horrifying: 66 people killed, more than a half trapped in their cars, hundreds of houses destroyed, more than 50 000 hectars of land burned and more than 50 milions of damages to affected activities[21].

Wilfires annually menace Portugal because of several unfavorable conditions. It is a warm country fanned by strong Atlantic winds and climate change is already extending the “fire season” with summers more and more dry and hot. Furthermore, Portugal is one of the most forested European countries, but the majority of forests are privately owned and exploited for wood. Monocultural forests of Eucalyptus characterized the scorched area, a very profitable but also easy burning tree. Finally, ineffective fire prevention strategies, communication problems and lack of firefighting equipment contributed to worsen an already harsh situation[22].

After days of efforts the 1600 firefighters involved, assisted by more humid and less warm days, managed to extinguish the fire. There are not certainties on the origin of the fire yet, even if local authorities pointed toward natural causes. What remains is a scorched scar in the heart of the country that brought, in the words of the Prime Minister António Costa, *“a dimension of human tragedy that we cannot remember”*[23].

In the following, it is reported the table containing DID estimation results of Pedrógão Grande wildfire. Almost all DID coefficients are significant at their highest significance level.

[21] Portugal's wildfire that broke a community - BBC News
[22] Portugal wildfires: Why are they so deadly? - BBC News
[23] Portugal Fires Kill More Than 60, Including Drivers Trapped in Cars - The New York Times (nytimes.com)

	TA	OR	CO	ROA
time	-0.2953749***	-0.1330521***	-0.1319311***	-0.0321207***
	(0.0078396)	(0.0080706)	(0.0079306)	(0.0017050)
A	0.1159703***	0.0744475***	0.0510193***	0.0277385***
	(0.0044248)	(0.0046161)	(0.0045002)	(0.0009895)
did	-0.0342926***	-0.0350459***	-0.0249134**	-0.0084012***
	(0.0096171)	(0.0100488)	(0.0097753)	(0.0021724)
age	0.0342222***	0.0187459***	0.0184995***	0.0002224***
	(0.0001589)	(0.0001531)	(0.0001482)	(0.0000318)
N	8.97e+05	8.72e+05	8.74e+05	8.92e+05

*Table 3.7 – Regression estimators for relevant indipendente and control variables of Pedrógão Grande wildfire 2017. Standard errors are reported in brackets, N is the number of observations; * significant at p<0.1; ** significant at p<0.05; *** significant at p<0.01.*

All variables show solid negative DID coefficients. Differently from other case study events, also ROA presents a negative coefficient, suggesting a redistribution of ratio weights in favour of Total assets.

Discussion and further developments

This final section is meant to discuss and raise conclusions on results of this thesis. Special attention will be also dedicated to implications and further developments for the current literature on this topic.

The analysis of interaction between extreme weather event exposure and firm performance was performed through the construction of two panel datasets, one containing EWE information and the other including company level financial data from more than eight million European firms. The process was carried out employing several data sources (Orbis, EMDAT, Eurostat) and multiple softwares (Python, QGIS, Excel, Stata), resulting in the construction of two databases with a level of completeness and detail in this area of knowledge without public available comparisons, to the best of author's knowledge.

The potential of these two datasets makes them a relevant resource for future reserches in the field, opening to several opportunities still unexplored. Indeed, information from the financial database has already been exploited by the Climate Finance Observatory at the Politecnico di Milano and in the development of a thesis work at the Politecnoco di Torino.

As already described in the literature review, very few studies addressed the EWE impact on firm performance and all of them involved financial data only from publicly listed firms. For these reasons, the work of this thesis represents an element of absolute novelty in the field referring essentially to SMEs and thus covering a very important and representative group of European enterprises.

Results of the analysis raise some interesting findings in the field. First of all, estimates, regardless of event type, does not draw a unique trend for considered variables with an almost balanced partition between positive and negative effects. This consideration highlight the complexitity of the interaction between firm performance and EWE impacts, in line with some papers in literature (e.g. Altay & Ramirez, 2010), that is influenced by several aspects related to both the event (such as severity, type, duration, predictability and extension of affected area) and impacted subjects (such as industrial sector, supply chain position, managements decisions and firm vulnerability).

Nevertheless, at least for TA, OR and CO, mean values of estimates exhibit, in almost every case, a negative trend, generally led by developed negative tails in the coefficients distribution. On the contrary, ROA draws a mean positive trend of coefficients for all category of event but with far lower absolte values of effects.

Moving on to case studies, they show a very coherent behaviour in their estimation. They all represent a severe meteorological event with a tragically remarkable death toll and a total amount of damages documented by international news coverage. These characteristics reflected into an overall strongly negative trend of coefficients.

Starting from the findings of the current thesis, several further developments in the field could be addressed.

First of all, it may be useful to investigate further firm and event characteristics to highlight eventual different behaviours from EWE exposure. Peculiarities such as the industrial sector and the supply chain position, as already suggested by other studies in literature (e.g. Altay & Ramirez, 2010 and Huang, Kerstein & Wang, 2018) are just some examples of possible integrations for this kind of research.

Additionally, it may be interesting to further exploit the financial data by exploring responses of other variables. In fact, the current work tests just the most representative and essential variables for firm performance. As already explained, this choice was guided by the need of immediate indicators and limited by the novelty of the explorative approach carried out, but the completeness of financial data offers a series of compelling opportunities.

Focusing on potential improvements for future applications in the field, at least two limiting factors in the current thesis can be addressed.

The first regards the ten-years extension of data. This aspect was mainly due to data availability issues, but it could be considered a relatively small time period with respect to usual ones addressed in literature. Thus, further applications should consider the possibility of a time expansion of the panel data.

Another limiting aspect is related to the geographical location methodology. The matching between events and firm locations relied on NUTS3 regions, that for sure allowed a level of event assignment detail never achieved before with such a number of events in the same analysis, but on the other hand it may incur in misestimation of affected subjects. For instance, an event that affected just a limited group of cities, through this method, is assigned to the whole NUTS3 region where the cities are located. The assumption of NUTS3 level as the more detailed region assignment was also guided by the location method adopted by EMDAT that in more than one occasion seems general and undetailed. For these reasons, large opportunities of improvements for further applications reside in this particular aspect concerning the location definition and matching.

The final considerations made above determine the end of the current thesis work.

Appendix

Selected Journals

Category	Name	H index
F	Academy of Management Annals	51
F	Academy of Management Journal	283
F	Academy of Management Review	242
E, F, M	Accounting Review	133
E	American Economic Review	253
E	Annual Review of Economics	39
E, M	Annual Review of Financial Economics	22
F	Annual Review of Organizational Psychology and Organizational Behaviour	31
M	Auditing	63
E, F	Brookings Papers on Economic Activity	75
M	Contemporary Accounting Research	81
M	Critical Perspectives on Accounting	57
E	Econometrica	169
E	Economic Journal	143
E,F	Entrepreneurship: Theory and Practice	121
M	European Economic Review	116
M	Family Business Review	87
M	Finance and Stochastics	38
M	Games and Economic Behavior	84
E	International Economic Review	79
F	International Organization	133
E, F, M	Journal of Accounting and Economics	132
E, F, M	Journal of Accountig Research	118
M	Journal of Banking and Finance	135
F	Journal of Business Venturing	154
F	Journal of Conflict Resolution	94
E, F	Journal of Consumer Research	155
M	Journal of Corporate Finance	83
E	Journal of Econometrics	135
E	Journal of Economic Growth	76
E	Journal of Economic Literature	145
E	Journal of Economic Perspectives	172
E, F, M	Journal of Finance	264

M	Journal of Financial and Quantitative Analysis	101
M	Journal of Financial Econometrics	31
E, F, M	Journal of Financial Economics	223
M	Journal of Financial Intermediation	67
E, F	Journal of Human Resources	92
E, F	Journal of International Business Studies	168
M	Journal of International Economics	121
M	Journal of International Management	60
E, F	Journal of Labor Economics	94
F, M	Journal of Management	192
E, F	Journal of Marketing	218
E, F	Journal of Marketing Research	147
E, M	Journal of Monetary Economics	112
M	Journal of Money, Credit and Banking	95
F	Journal of Operations Management	166
E	Journal of Political Economy	168
F	Journal of Public Administration Research and Theory	94
M	Journal of Public Economics	123
M	Journal of Risk and Uncertainty	65
F	Journal of Supply Chain Management	79
E, F	Journal of the Academy of Marketing Science	148
M	Journal of World Business	95
M	Long Range Planning	89
M	Management Accounting Research	76
F	Management Science	221
F	Manufacturing and Service Operations Management	71
E, F	Marketing Science	113
M	Mathematical Finance	68
F	MIS Quarterly: Management Information Systems	195
P	Nature	1096
E	NBER Macroeconomics Annual	56
F	Organization Science	211
F	Personnel Psychology	124
P	Proceedings of the National Academy of Sciences of the United States of America	699
F	Public Administration Review	115
E	Quantitative Economics	20
E	Quarterly Journal of Economics	228
E	RAND Journal of Economics	98
M	Real Estate Economics	54
F	Research in Organizational Behavior	57
E	Review of Economic Dynamics	53

E	Review of Economic Studies	124
E	Review of Economics and Statistics	142
M	Review of Finance	47
E, F, M	Review of Financial Studies	157
P	Science	1058
M	SIAM Journal on Financial Mathematics	24
F	Strategic Management Journal	253
M	Surveys in Operations Research and Management Science	17
M	World Bank Economic Review	79

Appendix Table 1 – List of journals used for the "Source Filter" step of the SLR screening process, with the use of subject categories (P for the top three impact factor journals, M for the SCImago "Business, Management and Accounting" subject area, F for the SCImago "Finance" subject area and E for the SCImago "Economics and Econometrics" subject area) and H indexes.

SLR Final Pool

Title	Authors	Year	Source title
Climate risk: The price of drought	Huynh T.D., Nguyen T.H., Truong C.	2020	Journal of Corporate Finance
Government effectiveness and institutions as determinants of tropical cyclone mortality	Tennant E., Gilmore E.A.	2020	PNAS
Borrowers under water! Rare disasters, regional banks, and recovery lending	Koetter M., Noth F., Rehbein O.	2020	Journal of Financial Intermediation
Coastal wetlands reduce property damage during tropical cyclones	Sun F., Carson R.T.	2020	PNAS
Mangroves protect coastal economic activity from hurricanes	del Valle A., Eriksson M. et al.	2020	PNAS
The Spillover Effects of Hurricane Katrina on Corporate Bonds and the Choice between Bank and Bond Financing	Massa M., Zhang L.	2020	Journal of Financial and Quantitative Analysis
Climatic Stress, Internal Migration, and Syrian Civil War Onset	Ash K., Obradovich N.	2020	Journal of Conflict Resolution
Normalized US hurricane damage estimates using area of total destruction, 1900−2018	Grinsted A., Ditlevsen P., Christensen J.H.	2019	PNAS
The effects of uncertainty on market liquidity: Evidence from Hurricane Sandy	Rehse D., Riordan R., Rottke N., Zietz J.	2019	Journal of Financial Economics
In the Path of the Storm: Does Distress Risk Cause Industrial Firms to Risk-Shift?	Aretz K., Banerjee S., Pryshchepa O.	2019	Review of Finance
Mapping the effects of drought on child stunting	Cooper M.W., Brown M.E. et al.	2019	PNAS
Long-run consequences of exposure to natural disasters	Karbownik K., Wray A.	2019	Journal of Labor Economics
Mangroves shelter coastal economic activity from cyclones	Hochard J.P., Hamilton S., Barbier E.B.	2019	PNAS
Climate change and educational attainment in the global tropics	Randell H., Gray C.	2019	PNAS
The price impact of extreme weather in developing countries	Heinen A., Khadan J., Strobl E.	2019	Economic Journal
How do banks react to catastrophic events? Evidence from Hurricane Katrina	Schüwer U., Lambert C., Noth F.	2019	Review of Finance
Climate risks and market efficiency	Hong H., Li F.W., Xu J.	2019	Journal of Econometrics

Graduated Flood Risks and Property Prices in Galveston County	Atreya A., Czajkowski J.	2019	Real Estate Economics
Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston	Zhang W., Villarini G. et al.	2018	Nature
The impact of climate risk on firm performance and financing choices: An international comparison	Huang H.H., Kerstein J., Wang C.	2018	Journal of International Business Studies
Weather shocks, agriculture, and crime: Evidence from India	Blakeslee D.S., Fishman R.	2018	Journal of Human Resources
Relationship between season of birth, temperature exposure, and later life wellbeing	Isen A., Rossin-Slater M. et al.	2017	PNAS
Do managers overreact to salient risks? Evidence from hurricane strikes	Dessaint O., Matray A.	2017	Journal of Financial Economics
Comparative advantage, capital destruction, and hurricanes	Pelli M., Tschopp J.	2017	Journal of International Economics
Long-run Health Repercussions of Drought Shocks: Evidence from South African Homelands	Dinkelman T.	2017	Economic Journal
Estimating economic damage from climate change in the United States	Hsiang S., Kopp R., Jina A. et al.	2017	Science
Climate change damages to Alaska public infrastructure and the economics of proactive adaptation	Melvin A.M., Larsen P. et al.	2017	PNAS
Civil conflict sensitivity to growing-season drought	Von Uexkull N., Croicu M. et al.	2016	PNAS
Social and economic impacts of climate	Carleton T.A., Hsiang S.M.	2016	Science
Bondholder Concentration and Credit Risk: Evidence from a Natural Experiment	Manconi A., Massa M., Zhang L.	2016	Review of Finance
Drought and civil war in sub-saharan Africa	Couttenier M., Soubeyran R.	2014	Economic Journal
Weather shocks, sweet potatoes and peasant revolts in historical China	Jia R.	2014	Economic Journal
Flooding and liquidity on the bayou: The capitalization of flood risk into house Value and Ease-of-Sale	Turnbull G.K., Zahirovic-Herbert V., Mothorpe C.	2013	Real Estate Economics
Revolt on the Nile: Economic Shocks, Religion, and Political Power	Chaney E.	2013	Econometrica
The economic growth impact of hurricanes: Evidence from U.S. coastal counties	Strobl E.	2011	Review of Economics and Statistics

Impact of disasters on firms in different sectors: Implications for supply chains	Altay N., Ramirez A.	2010	Journal of Supply Chain Management
Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America	Hsiang S.M.	2010	PNAS
How disasters affect local labor markets: The effects of hurricanes in Florida	Belasen A.R., Polachek S.W.	2009	Journal of Human Resources
Valuing flood disasters using the life satisfaction approach	Luechinger S., Raschky P.A.	2009	Journal of Public Economics
Does the exchange rate regime matter for real shocks? Evidence from windstorms and earthquakes	Ramcharan R.	2007	Journal of International Economics
Europe-wide reduction in primary productivity caused by the heat and drought in 2003	Ciais Ph., Reichstein M. et al.	2005	Nature

Appendix Table 2 – List of 41 articles selected through the full review step to be included in the final SLR pool.

List of Figures

List of Tables

Acronyms

A: affected firms dummy variable

CO: logarithmic Costs of Operations

DID: Difference-in-Difference

EWE: Extreme Weather Event

mOR: arithmetical Mean of Operating Revenues in the three years before the event

mTA: arithmetical Mean of Total Assets in the three years before the event

OR: logarithmic Operating Revenues

PNAS: Proceedings of the National Academy of Sciences of the United States of America

PSM: Propensity Score Matching

ROA: Return Over Assets

SLR: Systematic Literature Review

SMEs: Small and medium-sized enterprises

TA: logarithmic Total Assets

TC: Tropical Cyclones

US: United States

Bibliography

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. Proceedings of the National Academy of Sciences of the United States of America, 113(42), 11770-11775.

Altay, N., & Ramirez, A. (2010). Impact of disasters on firms in different sectors: Implications for supply chains. Journal of Supply Chain Management, 46(4), 59-80.

Aretz, K., Banerjee, S., & Pryshchepa, O. (2019). In the path of the storm: Does distress risk cause industrial firms to risk-shift? Review of Finance, 23(6), 1115-1154.

Ash, K., & Obradovich, N. (2020). Climatic stress, internal migration, and syrian civil war onset. Journal of Conflict Resolution, 64(1), 3-31.

Atreya, A., & Czajkowski, J. (2019). Graduated flood risks and property prices in galveston county. Real Estate Economics, 47(3), 807-844.

Belasen, A. R., & Polachek, S. W. (2009). How disasters affect local labor markets: The effects of hurricanes in florida. Journal of Human Resources, 44(1), 251-276.

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, 327(5964), 454-458.

Blakeslee, D. S., & Fishman, R. (2018). Weather shocks, agriculture, and crime: Evidence from india. Journal of Human Resources, 53(3), 750-782.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., . . . Živković, N. (2019). Changing climate both increases and decreases european river floods. Nature, 573(7772), 108-111.

Carleton, T. A., & Hsiang, S. M. (2016). Social and economic impacts of climate. Science, 353(6304).

Chaney, E. (2013). Revolt on the nile: Economic shocks, religion, and political power. Econometrica, 81(5), 2033-2053.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., . . . Valentini, R. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. Nature, 437(7058), 529-533.

Cooper, M. W., Brown, M. E., Hochrainer-Stigler, S., Pflug, G., McCallum, I., Fritz, S., . . . Zvoleff, A. (2019). Mapping the effects of drought on child stunting. Proceedings of the National Academy of Sciences of the United States of America, 116(35), 17219-17224.

Couttenier, M., & Soubeyran, R. (2014). Drought and civil war in sub-saharan africa. Economic Journal, 124(575), 201-244.

Dankers, R., Arnell, N. W., Clark, D. B., Falloon, P. D., Fekete, B. M., Gosling, S. N., . . . Wisser, D. (2014). First look at changes in flood hazard in the inter-sectoral impact model intercomparison project ensemble. Proceedings of the National Academy of Sciences of the United States of America, 111(9), 3257-3261.

del Valle, A., Eriksson, M., Ishizawa, O. A., & Miranda, J. J. (2020). Mangroves protect coastal economic activity from hurricanes. Proceedings of the National Academy of Sciences of the United States of America, 117(1), 265-270. doi:10.1073/pnas.1911617116

Dessaint, O., & Matray, A. (2017). Do managers overreact to salient risks? evidence from hurricane strikes. Journal of Financial Economics, 126(1), 97-121.

Diffenbaugh, N. S., Swain, D. L., Touma, D., & Lubchenco, J. (2015). Anthropogenic warming has increased drought risk in california. Proceedings of the National Academy of Sciences of the United States of America, 112(13), 3931-3936.

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from south african homelands. Economic Journal, 127(604), 1906-1939. doi:10.1111/ecoj.12361

Duffy, P. B., Brando, P., Asner, G. P., & Field, C. B. (2015). Projections of future meteorological drought and wet periods in the amazon. Proceedings of the National Academy of Sciences of the United States of America, 112(43), 13172-13177.

Garner, A. J., Mann, M. E., Emanuel, K. A., Kopp, R. E., Lin, N., Alley, R. B., . . . Pollard, D. (2017). Impact of climate change on new york City's coastal flood hazard: Increasing flood heights from the preindustrial to 2300 CE. Proceedings of the National Academy of Sciences of the United States of America, 114(45), 11861-11866.

Grinsted, A., Ditlevsen, P., & Christensen, J. H. (2019). Normalized US hurricane damage estimates using area of total destruction, 1900−2018. Proceedings of the National Academy of Sciences of the United States of America, 116(48), 23942-23946.

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2012). Homogeneous record of atlantic hurricane surge threat since 1923. Proceedings of the National Academy of Sciences of the United States of America, 109(48), 19601-19605.

Hansen, J., Sato, M., & Ruedy, R. (2012). Perception of climate change. Proceedings of the National Academy of Sciences of the United States of America, 109(37), E2415-E2423.

Heinen, A., Khadan, J., & Strobl, E. (2019). The price impact of extreme weather in developing countries. Economic Journal, 129(619), 1327-1342.

Hochard, J. P., Hamilton, S., & Barbier, E. B. (2019). Mangroves shelter coastal economic activity from cyclones. Proceedings of the National Academy of Sciences of the United States of America, 116(25), 12232-12237.

Holden, Z. A., Swanson, A., Luce, C. H., Jolly, W. M., Maneta, M., Oyler, J. W., . . . Affleck, D. (2018). Decreasing fire season precipitation increased recent western US forest wildfire activity. Proceedings of the National Academy of Sciences of the United States of America, 115(36), E8349-E8357.

Hong, H., Li, F. W., & Xu, J. (2019). Climate risks and market efficiency. Journal of Econometrics, 208(1), 265-281.

Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., . . . Houser, T. (2017). Estimating economic damage from climate change in the united states. Science, 356(6345), 1362-1369.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. Proceedings of the National Academy of Sciences of the United States of America, 107(35), 15367-15372.

Huang, H. H., Kerstein, J., & Wang, C. (2018). The impact of climate risk on firm performance and financing choices: An international comparison. Journal of International Business Studies, 49(5), 633-656.

Huynh, T. D., Nguyen, T. H., & Truong, C. (2020). Climate risk: The price of drought. Journal of Corporate Finance, 65.

Isen, A., Rossin-Slater, M., Walker, R., & Smith, V. K. (2017). Relationship between season of birth, temperature exposure, and later life wellbeing. Proceedings of the National Academy of Sciences of the United States of America, 114(51), 13447-13452.

Jia, R. (2014). Weather shocks, sweet potatoes and peasant revolts in historical china. Economic Journal, 124(575), 92-118.

Karbownik, K., & Wray, A. (2019). Long-run consequences of exposure to natural disasters. Journal of Labor Economics, 37(3), 949-1007.

Koetter, M., Noth, F., & Rehbein, O. (2020). Borrowers under water! rare disasters, regional banks, and recovery lending. Journal of Financial Intermediation, 43.

Lai, Y., Li, J., Gu, X., Chen, Y. D., Kong, D., Gan, T. Y., . . . Wu, G. (2020). Greater flood risks in response to slowdown of tropical cyclones over the coast of china. Proceedings of the National Academy of Sciences of the United States of America, 117(26), 14751-14755.

Luechinger, S., & Raschky, P. A. (2009). Valuing flood disasters using the life satisfaction approach. Journal of Public Economics, 93(3-4), 620-633.

Manconi, A., Massa, M., & Zhang, L. (2016). Bondholder concentration and credit risk: Evidence from a natural experiment. Review of Finance, 20(1), 127-159.

Marlon, J. R., Bartlein, P. J., Gavin, D. G., Long, C. J., Anderson, R. S., Briles, C. E., . . . Walsh, M. K. (2012). Long-term perspective on wildfires in the western USA. Proceedings of the National Academy of Sciences of the United States of America, 109(9), E535-E543.

Massa, M., & Zhang, L. (2020). The spillover effects of hurricane katrina on corporate bonds and the choice between bank and bond financing. Journal of Financial and Quantitative Analysis.

Mazdiyasni, O., & AghaKouchak, A. (2015). Substantial increase in concurrent droughts and heatwaves in the united states. Proceedings of the National Academy of Sciences of the United States of America, 112(37), 11484-11489.

Meehl, G. A., & Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. Science, 305(5686), 994-997.

Melvin, A. M., Larsen, P., Boehlert, B., Neumann, J. E., Chinowsky, P., Espinet, X., Marchenko, S. S. (2017). Climate change damages to alaska public infrastructure and the economics of proactive adaptation. Proceedings of the National Academy of Sciences of the United States of America, 114(2).

Neelin, J. D., Sahany, S., Stechmann, S. N., & Bernstein, D. N. (2017). Global warming precipitation accumulation increases above the current-climate cutoff scale. Proceedings of the National Academy of Sciences of the United States of America, 114(6), 1258-1263.

Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G. J., Allen, M. R. (2011). Anthropogenic greenhouse gas contribution to flood risk in england and wales in autumn 2000. Nature, 470(7334), 382-385.

Pelli, M., & Tschopp, J. (2017). Comparative advantage, capital destruction, and hurricanes. Journal of International Economics, 108.

Ramcharan, R. (2007). Does the exchange rate regime matter for real shocks? evidence from windstorms and earthquakes. Journal of International Economics.

Randell, H., & Gray, C. (2019). Climate change and educational attainment in the global tropics. Proceedings of the National Academy of Sciences of the United States of America, 116(18), 8840-8845.

Reed, A. J., Mann, M. E., Emanuel, K. A., Lin, N., Horton, B. P., Kemp, A. C., & Donnelly, J. P. (2015). Increased threat of tropical cyclones and coastal flooding to new york city during the anthropogenic era. Proceedings of the National Academy of Sciences of the United States of America, 112(41).

Rehse, D., Riordan, R., Rottke, N., & Zietz, J. (2019). The effects of uncertainty on market liquidity: Evidence from hurricane sandy. Journal of Financial Economics.

Runyan, R.C. (2006) 'Small business in the face of crisis: identifying barriers to recovery from a natural disaster'. Journal of Contingencies and Crisis Management. 14(1). pp. 12–26.

Schüwer, U., Lambert, C., & Noth, F. (2019). How do banks react to catastrophic events? evidence from hurricane katrina. Review of Finance, 23(1), 75-116. doi:10.1093/rof/rfy010
Sheffield, J., Wood, E. F., & Roderick, M. L. (2012). Little change in global drought over the past 60 years. Nature.

Silvestro, F, Rebora, N, Giannoni, F, Cavallo, A, Ferraris, L (2016) 'The flash flood of the Bisagno Creek on 9th October 2014: An ''unfortunate" combination of spatial and temporal scales' Journal of Hidrology Vol. 541 (part A) pp. 50-62.

Strobl, E. (2011). The economic growth impact of hurricanes: Evidence from U.S. coastal counties. Review of Economics and Statistics.

Sun, F., & Carson, R. T. (2020). Coastal wetlands reduce property damage during tropical cyclones. Proceedings of the National Academy of Sciences of the United States of America, 117(11).

Tennant, E., & Gilmore, E. A. (2020). Government effectiveness and institutions as determinants of tropical cyclone mortality. Proceedings of the National Academy of Sciences of the United States of America, 117(46).

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3).

Trapp, R. J., Diffenbaugh, N. S., Brooks, H. E., Baldwin, M. E., Robinson, E. D., & Pal, J. S. (2007). Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. Proceedings of the National Academy of Sciences of the United States of America, 104(50).

Trouet, V., Harley, G. L., & Domínguez-Delmás, M. (2016). Shipwreck rates reveal caribbean tropical cyclone response to past radiative forcing. Proceedings of the National Academy of Sciences of the United States of America, 113(12).

Turnbull, G. K., Zahirovic-Herbert, V., & Mothorpe, C. (2013). Flooding and liquidity on the bayou: The capitalization of flood risk into house value and ease-of-sale. Real Estate Economics, 41(1), 103-129.

Von Uexkull, N., Croicu, M., Fjelde, H., & Buhaug, H. (2016). Civil conflict sensitivity to growing-season drought. Proceedings of the National Academy of Sciences of the United States of America, 113(44), 12391-12396.

Zhang, W., Villarini, G., Vecchi, G. A., & Smith, J. A. (2018). Urbanization exacerbated the rainfall and flooding caused by hurricane harvey in houston. Nature, 563(7731), 384-388.