Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE STATISTICHE

Ciclo 34

Settore Concorsuale: 13/D4 - METODI MATEMATICI DELL'ECONOMIA E DELLE SCIENZE ATTUARIALI E FINANZIARIE

Settore Scientifico Disciplinare: SECS-S/06 - METODI MATEMATICI DELL'ECONOMIA E DELLE SCIENZE ATTUARIALI E FINANZIARIE

TOPICS IN CLIMATE CHANGE RISK

Presentata da: Leonardo Bortolan

Coordinatore Dottorato

Monica Chiogna

Supervisore

Silvia Romagnoli

Esame finale anno 2022

Contents

List of Figures

List of Tables

Introduction

This thesis encompasses three works that analyze climate change risk implications in financial instruments.

Specifically, we analyze how climate risk impacts economic players and its consequences on the financial markets. Essentially, literature unravels two main channels through which climate change poses risks to the status quo, namely physical and transitional risk, that we cover in three works.

Firstly, the call for a global shift to a net-zero economy, implicitly devalues assets contributing to global warming that regulators are forcing to dismiss.

On the other hand, abnormal changes in the temperatures as well as weather-related events challenge the environmental equilibrium and could directly affect firms' operations as well as profitability.

We start the analysis with the physical component, by presenting a statistical measure that generally represents shocks to the distribution of temperature anomalies. We oppose this statistic to classical physical measures and assess that it is the driver of the electricity consumption, in the weather derivatives market, and the cross-section of equity returns. We find two transmission channels, namely investor attention, and firm operations.

We then analyze the transition risk component, by associating a regulatory horizon characterization to fixed income valuation. We disentangle a risk driver for corporate bond overperformance that is tight to change in credit riskiness, and, after applying a statistical learning model to forecast excess returns, we include carbon emission metrics without clear evidence of its importance. Finally, we analyze the effect of

change in carbon emissions on a regulated market such as the EU Emission Trading Scheme (ETS) by selecting utility sector corporate bonds and, after controlling for possible risk factors, we document how a firm's carbon profile differently affects the term structure of credit riskiness.

The Thesis is composed of four parts.

The first chapter represents an introduction to climate change and the interplay with economic activities. We analyze the statistical properties of climate change and the related literature, which constitutes the baseline for the research questions, elicited in the last section of the chapter.

The second chapter represents the first working paper which deals with physical risk. We develop a statistical measure to generally represent extremes in the distribution of temperature anomalies that describes weather exposure and we contrast it to the most employed by researchers to assess its consequences on financial markets.

The third chapter is constituted by a work in which we disentangle a market invariant for corporate bond excess returns and, through a statistical learning model, we assess the projection at the horizon to forecast bond overperformance. In addition, we test whether transition risk variables are a driving component.

Finally, the last chapter analyses the transition risk horizon in the utility corporate bond market, revealing the component that drives this kind of risk.

Chapter 1

Literature on Climate Change Risk and Finance

1.1 Phisics of Climate change

The analysis regarding the effects climate change can pose to the various economic players should originate from the analysis of how climate change arises, from a climatological perspective. When considering financial instruments, it is possible to disentangle and price its future dynamic just when linked to a specific risk factor. This is an essential point because in this thesis we crosscut different sciences, namely statistics, economics, and climatology.

The first work involving the interaction of climate and economics is dated back to 1977 by Nordhaus (Nordhaus (1977)).

It is important to tackle the problem firstly by the physical side, that is the underlying component of climatic models, to have a better clue, in the proceedings of this work, on the assumptions and the terminology employed. Hsiang and Kopp (2018) assess that physical sciences are a minor in economic research, which results in lower interest for a crucial risk driver for the models that are being developed.

When analyzing the temperature equilibrium level the planet has reached in this era, we start by considering the energy the planet incorporates that is determined by two

main factors:

- Sunlight Energy: energy coming in the visible and ultraviolet space
- Earth Radiation: energy in the infrared space that is backed by the earth to the space.

So, if one of these two components changes, the equilibrium temperature changes as well. In addition, it is necessary to express the two components as a function of many other factors. The sunlight energy depends, among other determinants, on the Sun's temperature, the distance Sun-Earth and Earth's reflectivity, a phenomenon known as albedo. It is sufficient that one of these variables slightly changes to experience an increase or decrease in the world equilibrium temperature. The most cited example to this point considers the albedo, which is determined by particular clouds that prevent 30% of the Sun's energy from even reaching the Earth. This value can increase with sulfurs emitted by volcanoes: around 536B.C. the average temperatures in the Nordic hemisphere declined by 1.5/2.5°C caused by the eruption of an Iceland's volcano (Toohey et al. (2019)).

On the other side, the earth radiates back to space some heat, represented by the amount of energy the Earth can eject back. The light space of the incoming radiation is different from the outgoing ones: incoming light is in the visible and ultraviolet region whereas outgoing is placed in the infrared. At this stage, it is possible to face greenhouse gases (not limited only to CO_2) they do not interact with ultraviolet sunlight but they do with infrared, absorbing a percentage of outgoing radiation that is sent back to the earth. However they are not only a negative component considering that, without this interaction, the equilibrium temperature at the earth level would also be around -18°C, solely determined by the sunlight energy.

The amount of greenhouse gases in the atmosphere determines the height from which infrared light can effectively escape Earth. It is possible to locate this height considering where the temperature is at -18°C, which in recent history has reached equilibrium at 5500m. As well explained (Hsiang and Kopp (2018), Hansen et al. (1981)),

a marginal increase in greenhouse gasses in the atmosphere produces a direct and an indirect effect. The direct effect is represented by the increase in the effective radiating level derived from the decrease in the total energy that can escape the atmosphere. The indirect effect instead is derived by the composition of the triggered effects the increase of temperature brings, with the most important being the water vapor, which is more present in the context of a warmer climate and absorbs the infrared radiation, thus amplifying the greenhouse direct effect.

For example (Hansen et al. (1981)), doubling CO_2 concentration theoretically would increase the average surface temperature by 1.2°C, but taking into consideration all the effects, the equilibrium point could vary between 2.0°C and 4.5°C. This result pushes us to consider two elements.

Firstly, the difficulty that climate models still face in generating baseline climate projection: even accessing high computational power, it is difficult to generate accurate predictive scenarios, given the enormous amount of interconnections between variables. The result is then a series of plausible scenarios analyzing a subfield of the parameter space, that is released by the Intergovernal Panel for Climate Change (IPCC).

Secondly, there is high uncertainty around the effect greenhouse gases produces. An example is extreme events, in which the number is increasing (Cai et al. (2014)) but there is no consensus around how to forecast their future occurrence given by global warming.

1.2 Statistical properties of climate and weather

We tackle the modeling framework of climate and weather from a statistical point of view, following the conjecture by Hsiang (2016) where the determinants of the climate conditions can be thought of as a random vector determined both in time, t, and space, s:

$$w_{s,t} = [temperature_{s,t}, precipitation_{s,t}, humidity_{s,t}, wind_{s,t}, ...] \tag{1.1}$$

Two important considerations arise. The first one is related to the fact that these variables influence not only the atmosphere but also related factors, like oceans, rivers, the glacier that can trigger important consequences.

The second analysis is related to the impact these variables have on socio-economic activities. Climate change is a threat just when it affects socially or economically the population. One can argue that this is myopic, but if an event, through the concatenation effect, doesn't impact in any way the socio-economic activities, either in the long or in the short run, then it is not a risk factor(Diaz and Moore (2017)).

The next step is to understand the fundamental difference between climate and weather, and, following the notation by Hsiang (2016), we employ a statistical approach.

Considering an open time interval $\tau = [t, t + \delta)$, the random vector representing the weather conditions comes from the probability distribution $\psi(C_{i\tau})$:

$$w_{st} \in \psi(C_i\tau), \quad \forall t \in \tau \tag{1.2}$$

with $C_{s\tau}$ being a vector representing the fundamental parameters of the family distribution ψ which results in the weather outcomes w_{it}. Weather conditions represent the random variables to observe and measure, a realization of the hidden true population, climate. The weather thus is the realization through which we can assess the changing behavior of the unknown process.

Finally, an important component is the data collection technique that arises in the context of meteorological data, often not considered in financial literature. It is possible to characterize meteorological data as continuous stochastic processes defined over a three-dimensional space (temperature, for example, can be measured either at land-level or at a certain constant height from sea-level) and then they need to

be transformed in a discrete space as well in the time. Weather data are primarily recorded from meteorological stations, a process that presents a main drawback when there are no stations for a particular area. To circumvent this problem, it is possible to infer measurements from spotted recordings to obtain a dataset also for geographical points not covered by stations. For the sake of the research, we employed both approaches and, we address the main concern regarding temperature data, namely Gap Filling, Value and Jump detection, following Jewson and Brix (2005). When dealing with Gap filling at hourly data, it is possible to compute daily maximum, minimum, or Diurnal Temperature Range (DTR, maximum minus minimum) even with a missing value. When a daily temperature data is missing it is possible to retrieve through the average of the previous and the following day[1].

$$T_d = \frac{T_{d-1} + T_{d+1}}{2} \tag{1.3}$$

In the case there are nearby stations available, however, the preferred method is a spatial interpolation procedure, by averaging data from different stations. Analyzing neighboring stations allows detecting abnormal values when temperatures differ considerably and appropriately adjusting the anomalous one. Finally, we take into account also jump detection. The main source of jumps in the time series is equipment replacement that are however easy to identify. More challenging is the variation in the surrounding area of the station, which causes changes in the measurement, such as urbanization, which is the main source of misleading information. Evidences (Karl et al. (1988), Ren et al. (2008)) suggest that urbanized area face an higher increase in temperature mean levels compared to rural area. A slow-moving increase of recorded temperature in a certain meteorological station can thus derive from its citified surrounding area rather than climate change.

[1]Alternative approaches, such as estimation of AR(p) / ARMA(p,q) models and then forward fill the missing value, are suggested by literature (Boissonnade et al. (2002)). The different dataset we employ presents minor missing data, in the order of a couple in fifty years.

1.3 Temperature and Weather Derivatives

The core component of the random vector representing climate – as defined in Equation (1.1)– is the temperature variable and was the first one incorporated in the literature following the introduction of a derivative market for weather-related variable employed as a hedging factor.

The documented repercussion of temperatures on electricity consumption (Bentzen and Engsted (1993)) and agricultural output (Adams et al. (1990)), boosted the attention to the possibility to set up a market where the participant could settle their relative position on climatic variables.

Weather derivatives were introduced by Chicago Mercantile Exchange(CME) in 1996 as futures and options contracts, written on a variety of weather events such as temperatures, wind, or rain. For example, through a temperature derivative, energy companies were able to transfer volumetric risk caused by unfavorable weather conditions from electricity price to weather variables(Jewson and Brix (2005)).

We analyze in more detail temperature-related derivatives, with the two main employed being the ones written on the indexes

- **Heating Degree Days**(HDD), in which the holder of the contract receives a payoff according to an index x that accounts for the days and the extent at which the temperatures negatively exceed a certain threshold in a fixed number of days (N_d, usually a calendar month), according to the following formula:

$$x_{HDD} = \sum_{i=1}^{N_d} max(T_0 - T_i, 0) \tag{1.4}$$

  where T_i is the average temperature for the day i

- **Cooling Degree Days**(CDD) works in the same way with the exception that the payoff is referred to the exceeding temperatures:

$$x_{CDD} = \sum_{i=1}^{N_d} max(T_i - T_0, 0) \tag{1.5}$$

For CME contract, T_0 is set at 65°F that corresponds to 18°C. Given that CDD and HDD indexes are mutually exclusive, it follows that in the northern hemisphere during atmospheric summer months, the HDD index doesn't have a positive value. The same applies in the winter months for the CDD index.

With the first contract being exchanged at the CME, different players entered the market contributing to its improvement and the research quest arise in the modeling of the asset underlying (temperature, wind, or rain) and the pricing of such derivatives.

The quest for modeling and correctly pricing such contracts experienced all the steps already encountered in the financial sector. The peculiarity of such contracts is that the underlying is a nontradable asset thus hindering the hedging portfolio approach proposed by Black and Scholes (1973). The only available procedure is the direct modeling of the underlying temperature, and literature built upon this.

We start by analysing the pioneer work related to the stochastic modelling of temperatures, following the works by Dornier and Queruel (2000), Benth and Šaltytė-Benth (2005) and Benth and Benth (2007).

We consider a complete probability space defined by the Wiener triplet $(\Omega, \mathcal{F}, P)$ with filtration $\{\mathcal{F}_t\}_t \geq 0$, and a Brownian motion B_t. The baseline model well employed in finance is the mean reverting Ornstein-Uhlenbeck(OU) process that characterizes the evolution of a certain variable, x, through the Stochastic Differential Equation(SDE):

$$dx_t = \theta(\mu - x_t)dt + \sigma dB_t \tag{1.6}$$

that is characterized by the stationary long term equilibrium level μ.

In the case of temperature, T_t, the underlying process display a trend and a seasonality component in the level, that depends on the geographical position. The natural expansion for OU process (1.6) proposed by Benth and Šaltytė-Benth (2005)) investigates the seasonal component:

$$dT_t = ds(t) - k(T(t) - s(t))dt + \sigma dB(t) \tag{1.7}$$

where s_t is a function that accounts for both trend and seasonality. Among the several ways in which it is possible to extract this two components(Jewson and Brix (2005)), the choice by Benth and Šaltytė-Benth (2005) is the Truncated Fourier Series, and thus $s(t)$ can be expressed as:

$$s(t) = a + bt + \sum_{i=1}^{I_1} a_i sin(2i\pi(t - f_i)/365) + \sum_{j=1}^{J_1} b_j cos(2j\pi(t - g_j)/365) \tag{1.8}$$

This characterization has been documented as highly accurate in describing the salient stylized characteristic of the temperature levels across a year.
The conditional variance implied by the model is however constant along the different months, divergent from the empirical evidence by Campbell and Diebold (2005) of a higher volatility in the winter months. For this reason, Benth and Benth (2007) updated model (1.7) by considering a time-varying volatility coefficient for T_t. Again, the choice for the modelling of this component is a Truncated Fourier series of the following form:

$$\sigma^2(t) = c + \sum_{i=1}^{I} c_j sin(2i\pi t/365) + \sum_{j=1}^{J} d_j cos(2j\pi t/365) \tag{1.9}$$

The fact that they analyze and model σ^2 instead of directly σ derives from the options pricing formula, where the former is met more often, and thus it simplifies pricing equations.
It is important to note that, as opposed to the model that governs the trend and seasonality for the mean, in the equation for the variance it is not present any term that governs the changing behavior. It is then assumed that the variance is stable over time and that climate change doesn't affect at all this measure.

A second alternative compared to a stochastic process, Campbell and Diebold (2005) propose a simple time-series approach with conditional mean and variance, to allow more simplicity in the derivation of derivatives pricing. Their proposed model

is:

$$T_t = Trend_t + Seasonal_t + \sum_{l=1}^{l} \rho_{t-l} T_{t-l} + \sigma_t \epsilon_r \tag{1.10}$$

where the seasonal volatility is modeled through a generalized autoregressive conditional heteroskedasticity process (GARCH, Bollerslev (1986)). In these settings, however, there is a miss-match between the left-hand and the right-hand side of the equation. The deseasonalized temperatures are run with respect to the temperature level, since the seasonal function acts as a mere component, without affecting the regression estimation.

1.4 Physical Risk

In the previous section, we discuss how to model temperatures, without considering the damage effects they produce. Temperature is not the only consequence of global warming nor the only driver of climate consequences, however it is a good proxy for major economic impacts. A popular way to translate climate change into economic damage is through Integrated Assessment Models (IAM) by Nordhaus (1977). Evolving from the first work, three main IAM models have been developed, namely the Dynamic Integrated model of Climate and the Economy (DICE, Nordhaus and Sztorc (2013)), the Climate Framework for Uncertainty, Negotiation and Distribution (FUND, Tol and Anthoff (2014)), and the Policy Analysis of Greenhouse Effect (PAGE, Hope (2011)). To seek simplicity, we focus on the different damage functions that the three models employ, that derives from the change in the average temperature with respect a pre-warming period, ΔT.

The DICE model exhibits a quadratic damage function that can be expressed through the equation:

$$D = \delta_a \Delta T + \delta_b \Delta T^2 \tag{1.11}$$

The damage function of FUND model presents two innovations. Firstly, the impact is produced by a function of temperature anomalies, $g(\Delta T^x)$ that can take into account non-linearity in the case $x > 1$. Secondly, it embraces the elasticity, ϵ, of change in

income due to climate change:

$$D = g(\Delta T^x) + \frac{Y_t}{Y_0}^{-\epsilon} \tag{1.12}$$

Finally, the PAGE model takes into account adaptation that acts in a twofold direction. It acts negatively as a cost (C_{adp}) that allows to obtain a positive effected by decreasing the total variability of temperature that produce a damage, T_{adp}.

$$D = \delta\Delta(T - T_{adp})^x + C_{adp} \tag{1.13}$$

The differences in the damage functions regard mainly the components that determine the negative effectd. While DICE model takes into account just effects related to changes in temperatures, FUND relates the temperature damage to the income in a particular region and PAGE takes into account also the adaptation. The assumption on the temperature is however identical. The driver is represented by shifts in average temperatures, ΔT, that can be amplified by a certain power effect.

IAM models present undoubted advantages, as the possibility to directly assess future economics consequences of climate change at the level of interest. (see eg. Calvin et al. (2013), Bansal et al. (2017), Hassler and Krusell (2012)) The high complexity of such models however comes at a certain cost (Diaz and Moore (2017)). In this brief review, we focus on the assumptions that the damage functions, D, are not fully able to capture. Regarding the economic damage, the models produce a shock that affects either total factor productivity (TFP) or the whole economic output. However, in the standard definition as in the DICE models, the economic growth is exogenous and then, damage at time t does not affect growth at time $t + 1$(Dietz and Stern (2015)).In contrast, empirical evidence suggests that also growth rate are affected by weather realization, especially in poor countries (Dell et al. (2012)).

Another critique of the damage functions regards the lack of capturing the parametric uncertainty around the temperature drivers. In this light, just PAGE and FUND models account for uncertainty, through Monte Carlo simulations. As well

documented (Neumann and Strzepek (2014), Cai et al. (2015)) IAM models fail to consider adequately the cost associated with tipping point or the event on extreme realizations. Related to these two points, the damage functions employed lack a solid empirical background (Moore and Diaz (2015)).

To empirically assess how weather-related risk – temperature fluctuations or rainfall events – impacts on the economics, it is possible to consider macro indicators such as gross domestic product growth(GDP), Total Factor Productivity (TFP) as well as micro indicators such as labor supply or productivity. At the micro-level, extremely hot temperatures reduce the number of working hours especially for those sectors that are human capital intensive and display a lower level of adaptation (Graff Zivin and Neidell (2014)). In addition, evidence suggests that the response to temperature anomalies exhibits a non-linear response: an increase of 1 degree produces a more severe outcome as higher is the starting point (Burke et al. (2015)). Rich countries are less exposed to climate change given a quicker adaptation process (Fankhauser and McDermott (2014)), but the non-linear effects reported by Burke affect them as well. When aggregating at the macro level, results are less evident. While no damage is found at the macroeconomic output by Burke et al. (2015), output levels, as well as growth rates, are affected especially when considering poor countries (Dell et al. (2012)).

Climate change in addition produces winners and losers. Looking at the agricultural sector, Hong et al. (2012) find that a long-short strategy based on firm location and a drought index, finds positive alpha, meaning that places with a higher drought risk are more exposed. When the analysis spans the entire U.S., Deschênes and Greenstone (2007) find that climate change increases annual profit by 2/4%. The drivers are the beneficial effects coming from the warming process that outrun negative ones. Finally, given the non-linearity in the response, analysis of the U.K. economy from the pre-industrial era by Donadelli et al. (2019) highlights that years characterized by higher temperature volatility are associated with a shock in the TFP. In this light, volatility impacts also as a welfare cost, defined as compensation to households for

not bearing the risk. From a different angle, Kotz et al. (2021) builds from the acknowledgment that annual average temperature doesn't capture the full spectrum of the temperatures distribution. They show that a nominal increase in temperature variability affects especially regions at mid-latitudes or the coastal ones.

Finally, looking at the channel from the micro to the macro level, Benson and Clay (2004) find the destruction of capital stock as a possible driver. In this context, the dissection of players that drive economic output should emphasize the empirical observed effect, which is represented by firms. The research around physical implications focuses on the association between firm exposure to extreme values and the subsequent financial implications. Addoum et al. (2020) find a bi-directional movements in earning for firms experiencing extreme temperatures, that means extremes temperatures don't represent a negative shock for all sectors. These extremes can be represented either by heatwaves, measured as amount of time spent above a certain temperature threshold, or by cold snaps, measured as the amount of time experienced below a critical temperature level. This is a crucial point linked to the non-linearity effect found in the macro-economic works: the threshold of temperature anomalies is a key component to determine the occurrence of extremes. Following a similar approach, Pankratz and Schiller (2021) documents the repercussion such events produce in the supply chain. Extreme weather conditions not only affect the exposed entity but also its customers. As a reaction, the customer is more likely to change its supplier after an extreme weather event. From a slightly different perspective, Kumar et al. (2019) computes the return sensitivity of stock to abnormal temperatures and assesses that the highest exposed exhibits lower returns.

1.5 Transition Risk

Transition risk arise as the push to diminish possible future physical damages. The transition to a net-zero economy implies dismissing certain productivity elements – such as coal or oil – that become stranded assets. As in the physical dimension, literature undergo two different methodological analysis. The first stream analyses the

cost of emitting additional CO_2 in the atmosphere. Alternatively, one can address whether the employment of polluting technologies by firm, is priced in the entity dynamics, either as loss in income or a liability burden.

When looking at the first approach, carbon should be priced as the polluting components that cause global warming. In this regard, the Social Cost of Carbon (SCC, Tol (1999)), is the discounted price of a marginal damage derived by the marginal GhG emission. In order to correctly estimate this value, IAM models are usually employed given the long term projection characteristics they exhibits. Different estimation have been made, starting from Nordhaus (2014) that employs the DICE model (1.11) to determine a price of $18.6.

1.6 Open Questions

We conclude this first chapter by highlighting what we consider to be relevant open questions that we develop in the following works.

Regarding physical risk, to understand and describe events that directly affect productive sectors, the attention is on temperature extreme events. This definition implies the computation of time spent above certain thresholds, that can be fixed or time-varying. Literature focusing solely on the drift of temperature anomalies is not able to capture effective damages brought by a multifaceted type of risk. We question whether it is possible to derive a broader statistical measure related to temperature anomalies that can capture in an easier and more general way the multiple effects linked to weather risk. In addition, the transmission channel that enables a weather-related event to affect a firm's productivity or customer behavior is to be directly assessed. Regarding transition risk, our question regards its implication in credit risk that affects institutional portfolios, by analyzing fixed income instruments. Firstly, we need to isolate a risk driver for corporate bond overperformance, that is solely able to explain future excess return. Then we ask whether classical transition risk approaches can forecast changes in credit riskiness.

Finally, we ask whether it is possible to associate the transition risk profile with the

term structure of credit riskiness. Transition risk could arise as a regulatory risk and, according to the definition, the associated exposure could differ according to the instrument payoff.

Chapter 2

Nailing down temperature volatility: analizying their impact on asset prices

Work in collaboration with Atreya Dey [1] and Luca Taschini [2]

2.1 Introduction

Financial markets respond to temperature extremes, but not all shocks are equal. The equilibrium price of an asset can be affected by changes in temperature due to an adjustment in investors' beliefs (Choi et al. (2020)) or directly through firm-level exposure to temperature (Addoum et al. (2021)). Others find a tenuous relationship between temperature fluctuations and financial outcomes (Addoum et al. (2020)). According to the scientific evidence, the probability distribution of temperature anomalies has increased by more than one standard deviation due to climate change since the 1950s (Hansen et al. (2012)). Additionally, the distribution has broadened, leading to a higher number of temperature extremes experienced glob-

[1]University of Edinburgh Business School, atreya.dey@ed.ac.uk
[2]Grantham Research Institute and ESRC Centre for Climate Change Economics and Policy *London School of Economics*, l.taschini1@lse.ac.uk

ally. Identifying a relevant representation of changes in temperature distributions is thus an important aspect of fully understanding changes in temperature extremes, and their associated impacts on financial and economic outcomes.

To characterize extremes in temperature, we create a spatially dependent statistic that reflects deviations in temperature variability from its historical mean, on a rolling basis. The measure, $TD\text{-}VAR$, describes the unconditional distribution of temperature, which allows for the general description of spatial changes in temperature extremes over time. The generality in characterizing extremes with $TD\text{-}VAR$ means that we are able to treat cold spells as equally harmful to economic activity as heatwaves. Additionally, we incorporate large fluctuations in day-to-day temperature volatility with the same reasoning. Crucially, through a battery of validation exercises and asset pricing tests, we assert that changes to historical temperature variability is a primary driver of financial markets. We continue our analysis by explicitly disentangling *how* $TD\text{-}VAR$ affects markets by investigating whether the shock directly impacts investors' concerns about climate change or firms.

Throughout the chapter, we contrast our methodology with existing literature that highlights abnormal temperatures (Addoum et al. (2020)), defined as the number of days in which temperatures exceed a threshold, sustained over a certain time period, causing an extreme temperature event. We define this broadly as a *temperature anomaly* or TD. Addoum et al. (2020), Pankratz and Schiller (2021), and Choi et al. (2020) focus on one side of temperature anomalies – heatwaves – and are reliant on thresholds to define salient extremes. Others, such as Addoum et al. (2021), account for both cold and warm anomalies, yet still use a form of a priori thresholding. We construct $TD\text{-}VAR$ to reflect the volatility or variability in temperatures, which is known to affect crop yields (Wheeler et al. (2000), Ceglar et al. (2016)), human health and mortality (Zanobetti et al. (2011)), economic growth (Donadelli et al. (2017), Kotz et al. (2021)), and asset prices (Makridis (2018)).

We construct $\widetilde{TD}$, the realized monthly anomaly, to serve as a counterpoint to

TD-VAR. To build both metrics, we utilize multiple geospatial temperature data sets. The first covers city-level temperature data from the U.S. National Oceanic and Atmosphere Administration. The second is a gridded 1-degree latitude by 1-degree longitude daily temperature data set from Berkeley Earth Surface Temperatures (BEST). After aggregating the data to the U.S. state and country levels, we illustrate the differences between TD and TD-VAR with a set of realistic theoretical examples at the monthly and daily frequency. The two measures are then compared and contrasted in each step of our validation and asset pricing exercises, gradually leading to the conclusion that TD-VAR is the salient measure.

Our validation exercises begin by confirming that deviations in temperature variability are a primary driver of energy consumption and prices in weather derivatives. Following the logic that energy consumption is sensitive to deviations in temperature variability, we perform time-series analysis connecting the seasonality of energy demand to the two metrics. The analysis produces positive significant coefficients for TD-VAR on aggregate energy demand, especially in the residential and industrial sectors. In contrast, temperature anomalies only result in a significant positive coefficient for the commercial sector – attributable to the sector's steady energy demand. We continue to substantiate our claims by testing the relationship of the two statistics to the Chicago Mercantile Exchange (CME) weather futures market, which is intrinsically connected to energy consumption. The combined evidence strongly suggests that fluctuations in temperature variability from its historical mean is a first-order factor in the highly related energy and weather futures market.

Next, we test whether the differential exposure of firms to temperature shocks affects the valuation of their stock price. The primary source of variation in our two asset pricing tests are TD and TD-VAR for a firm headquartered in a U.S. state. Our first set of tests includes monthly cross-sectional return regressions to test the materiality of both temperature factors on Russell 3000 firms. The results indicate that both metrics are insignificant when considering firms in aggregate, corroborating the findings of Addoum et al. (2020). However, we find a positive significant coefficient

for $TD\text{-}VAR$ in the energy, utilities, consumer staples, and consumer discretionary sectors. The energy and utility sectors are again affected by temperature variability, reinforcing our validation exercises. A simple explanation for the consumer-related sectors is that shopping is difficult during cold spells or heatwaves, consistent with consumer demand and labor productivity channels (Starr (2000) and Griffin et al. (2017)) and documented in Donadelli et al. (2020), Colacito et al. (2019), Pankratz and Schiller (2021), and Addoum et al. (2021). Average monthly anomalies, on the other hand, are only economically consequential for the utilities sector. The return patterns are robust when we adjust the sample size to various sub-periods.

We continue to test whether investors are pricing deviations in temperature variability by implementing a dynamic investing strategy which goes long on firms headquartered in states that are least affected, and going short on those headquartered in the most-affected states. We build monthly quintile portfolios by ranking states on their exposure to the temperature metrics and place Russell 3000 firms in the differentially exposed portfolios based on their headquarter location. We rebalance the portfolios monthly, depending on the states' exposure to a temperature statistic. The methodology retrieves a time series of portfolio returns which we market-adjust with the Fama-French three factors and an additional momentum factor akin to Barber and Odean (2008). This strategy results in a 50% market-adjusted return over the 14-year sample period when using $TD\text{-}VAR$, compared to a negligible 8% (equivalent to 0.05% monthly) for temperature anomalies when considering all firms. Sectorally, the returns are markedly more for the energy and utilities sectors and are minor for the consumer sectors when using deviations in temperature variability. The results remain robust when we remove firms in states that are continually exposed to temperature shocks. The findings suggest that investors do account for temperature extremes by dynamically hedging the risks.

The aggregate results point to the fact that $TD\text{-}VAR$ is a salient physical risk for financial markets; however, the evidence we present is the combined effect of the material impact on firm operations and investor attention to climate change or

temperature. Temperature shocks have an impact on the attention of investors (Choi et al. (2020); Alekseev et al. (2021); Pastor et al. (2021)) and a corporation's earning processes (Addoum et al. (2021)), which together affect the equilibrium price of an asset. We attempt to explicitly disentangle the two effects by first associating the temperature metrics to *innovations* in attention indices, and second by extracting the *true* physical exposure of the firm, divorced from investor attention.

We investigate the attention channel at the U.S. country– and state level after aggregating the temperature metrics to the respective granularity. We adopt the innovations of *The Wall Street Journal* (WSJ) news index from Engle et al. (2020), which captures media coverage of climate change tailored to investors. This index only captures investor attention indirectly, as investors may not necessarily read the news, but we expect that temperature shocks would alert the media and investors to the negative implications of climate change. Additionally, we acquire the innovations of Google Search Volume Index (SVI) data at the state level for the topics "Climate Change" and "Temperature". Using this data, we explore whether retail investors react to localized temperature shocks (Alekseev et al. (2021)). Our U.S.-level results show a moderately significant relationship between $TD\text{-}VAR$ and unexpected news in the WSJ index. Similarly, there is a strongly significant relationship between our metric and both search topics. The only discernible relationship for temperature anomalies in both exercises is a positive significant relationship to searches for "Temperature". Collectively, the results suggest that shocks to temperature variability act as a "wake-up call" for investors.

The final exercise is to measure the attention-adjusted firm-level impact of temperature shocks. Due to data constraints and for reasons of simplicity, we assume that a firm's operations will be affected in the same state as their headquarters. If a temperature shock is truly salient to the processes of a firm, analysts will raise the issue of physical risk during earnings conference calls. We obtain a measure of firm-level exposure to physical risks from Sautner et al. (2020) who capture the pertinent discussion from earnings calls. Their measure, however, is influenced by attention to

climate change as seen by their positive association with the WSJ index. To disentangle this effect, we obtain residuals from a regression of expected and unexpected news from the WSJ on their physical exposure series. The residuals, which we call NetExposure, can be considered the 'true' impact of physical risks on a firm. We find that elevated $TD\text{-}VAR$ is significantly positively associated with an increase in NetExposure of a firm in comparison with $\widetilde{TD}$. Additionally, we find much larger effect sizes for the utilities and energy sector with our metric. Similar to our prior exercises, we find a muted effect on consumer sectors. In total, the results contextualize the price reaction by showing that $TD\text{-}VAR$ is salient for investors and firm operations.

Our primary contribution is to the nascent research on the financial consequences of temperature. Most work on climate shocks and financial markets defines abnormal temperatures as temperature extremes, i.e. temperatures being above a certain threshold, and investigates its effects on the international food industry (Hong et al. (2020)), firm earnings or profits (Pankratz and Schiller (2021); Addoum et al. (2020)), futures markets (Schlenker and Taylor (2021)), or asset prices in general (Bansal et al. (2017)). Addoum et al. (2021) incorporates the effects of both cold spells and heatwaves on industry earnings. One more relevant paper is that of Donadelli et al. (2017) who primarily investigate the effect of temperature volatility on macroeconomic outcomes and also find that shocks diminish U.K. and European equity prices in the cross-section.

We extend this literature in a number of ways. Our metric describes changes in the distribution of temperature anomalies, which reflects extremes more comprehensively than thresholds do. We document that deviations in temperature variability can represent extreme temperatures in a way that is salient for both weather futures and stock markets. While Addoum et al. (2021) focus on earnings, our results are in line with their findings that markets react both to extremely hot and extremely cold days. Furthermore, we find that day-to-day swings in temperature are similarly

consequential for the same markets. Our study also expands the equity pricing tests of anomalies in temperature volatility of Donadelli et al. (2017) in concrete manners: (1) by directly validating the importance of deviations in temperature variability in the energy consumption and weather futures market; (2) performing an asset pricing factor analysis to examine the relationship between deviations in temperature variability and U.S. stock returns; and (3) developing a long–short strategy to explore investor reactions to sub-national heterogeneity in temperature in the U.S. equity market.

Another contribution is to the body of research that studies the impact of temperature extremes on investor reactions and attention. Engle et al. (2020) builds the WSJ climate news series to hedge against long-term climate risks. Choi et al. (2020) finds that local temperature shocks can heighten investors' attention, which in turn differentially affects returns on a cross-section of stocks. Alekseev et al. (2021), with a similar argument, investigates the effects of local temperature shocks, finding that mutual funds respond by shifting their portfolio allocation, irrespective to the intensity of the heat shocks. We complement the majority of these findings, similarly concluding that investors do react to temperature swings. Specifically, deviations in temperature variability lead either to direct attention to a local shock, as in Choi et al. (2020) and Alekseev et al. (2021), or to indirect investor attention to increased news coverage distributed more broadly. Critically, however, we discover that the pricing reaction only occurs in response to a specific type of temperature shock. Furthermore, we go a step further by disentangling attention from firm-level exposure to the risk.

This chapter is organized as follows. Section 2.2 describes the data and explains our data set construction procedure. Section 2.3 describes how we expand upon prior temperature statistics and derive $TD\text{-}VAR$, the deviation in temperature variability. This section also illustrates that $TD\text{-}VAR$ comprehensively quantifies the extremes in the distribution of temperature. Section 2.4 validates $TD\text{-}VAR$ using electricity consumption and weather derivatives. The investigation into the relationships

between $TD\text{-}VAR$ and the U.S. equity market follows in Section 2.5. Section 2.6 disentangles the effect of $TD\text{-}VAR$ on investors' attention and isolates its direct impact on firms. Section 2.7 concludes.

2.2 Data

In this work, we consider data from a variety of sources. We start with geospatial temperature data that have a temporal characteristic. We link the spatial feature to financial variables by extrapolating their location. Finally, attention indexes are aggregated at national and sub-national levels.

2.2.1 Temperature data

Our analysis is based on the conjecture that deviations in temperature variability are more likely to affect the profitability of the U.S. corporate sector than are changes in mean temperature. The effects associated with deviations in temperature variability are likely to vary substantially across time and location.

Table 2.1. Specification of city dataset

City	GHCND Code	State	Weather Derivative	Mean	Std
Atlanta	GHCND:USW00013874	Georgia	X	72.0	15.3
Boston	GHCND:USW00014739	Massachusetts		59.3	18.4
Baltimore Washington	GHCND:USW00093721	Maryland		65.6	18.6
Cincinnati	GHCND:USW00093814	Ohio	X	63.8	19.7
Chicago	GHCND:USW00094846	Illinois	X	59.0	21.5
Dallas Forth Woot	GHCND:USW00093904	Texas		79.6	16.1
Des Moines	GHCND:USW00014933	Iowa		60.2	22.7
Detroit	GHCND:USW00014822	Michigan		58.7	20.7
Las Vegas	GHCND:USW00023169	Nevada	X	80.2	18.5
Minneapolis	GHCND:USW00014922	Minnesota	X	54.9	24.1
New York La Guardia	GHCND:USW00014732	New York	X	62.3	18.4
Portland	GHCND:USW00024229	Oregon	X	63.0	14.5
Philadelpia	GHCND:USW00013739	Pennsylvania		64.4	18.8
Salt Lake City	GHCND:USW00024127	Utah		64.3	21.2
Tucson	GHCND:USW00023160	Arizona		83.2	14.9

Table(2.1). Present the characteristics of the city for which we obtain GHNC daily data from NOAA. The stations coincide with the city airport. The columns weather derivatives indicates the stations for which a Cooling Degree Days (CDD) or Heating Degrees Day (HDD) weather derivative is traded at CME.

We collect data to test the relationship between deviations in temperature variability and weather derivatives. In particular, we obtain city-level temperature data for 13 U.S. cities[3] using the U.S. National Oceanic and Atmosphere Administration (NOAA) repository. Specifically, we use the NOAA Global Historical Climatology Network daily (GHNCd), an integrated database of climate summaries from land surface stations across the globe. For each city, we select the station corresponding to the closest city airport and retrieve the daily maximum temperature from the GHNCd.[4] The temperature data cover the period from 1 January 1950 to 31 December 2021. We pre-process the data by filling in missing values with the average of the maximum temperature recorded on the days either side. Table 2.1 reports the cities in our data set, the GHNCd code, the mean daily temperature, and the corresponding standard deviation.

The majority of our empirical analysis requires a spatially uniform, rather than city-specific, estimate of temperature. To measure location-specific exposure, we obtain spatially homogeneous daily temperature from the BEST database.[5] The BEST data are in the form of gridded 1-degree latitude by 1-degree longitude reconstruction of daily temperatures. BEST spatial interpolation provides extensive spatial coverage from 1950 to the present. BEST utilizes data from significantly more land stations (over 40,000) than the 10,000 used by alternative data sets, improving the assessment of record-setting daily U.S. temperatures. The data is then used to compute state-level and U.S.-wide daily temperatures, the details of which are presented in Section 2.3.4 and Appendix A.1.

[3]These are: Atlanta, ATL; Boston, BOS; Baltimore Washington, BWI; Chicago, ORD; Cincinnati, CVG; Dallas Fort Worth, DFW; Des Moines, DSM; Detroit, DTW; Las Vegas, LAS; Minneapolis St Paul, MSP; New York, LGA; Portland, PDX; Philadelphia, PHL; Salt Lake City, SLC, and Tucson, TUS. This same data was used in Diebold and Rudebusch (2022).

[4]Later in the analysis, we consider temperature-related weather derivative instruments. These contracts are city-specific and are settled based on the temperature readings of a specific weather station near the contract city.

[5]The application of homogenization techniques to daily temperature data is important in order to accurately understand the evolution of temperature extremes over the past century. The BEST daily temperature data use of a novel homogenized gridded approach to improve the assessment of record-setting daily U.S. temperatures. We refer to Rohde et al. (2013) and Rohde and Hausfather (2020) for a technical discussion.

2.2.2 Financial and economics data

We collect data on returns for the Russell 3000, an index tracking the performance of the 3,000 largest U.S. companies, representing approximately 97% of the investable U.S. equity market. Data on Russell 3000 constituents, their firm fundamentals and headquarter locations are drawn from Refinitiv, and they are classified into their respective sectors using the Global Industry Classification Standard (GICS).

Table 2.2. Summary Statistics, Russel 3000

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
LOGSIZE	9.18	9.21	9.08	8.97	9.10	9.17	9.18	9.28	9.34	9.34	9.32	9.40	9.43	9.42	9.40
B/M	12.88	14.38	15.03	13.69	13.99	14.55	14.45	14.99	15.69	16.11	15.69	15.91	16.75	17.14	17.82
ROE	11.25	9.04	-1.38	3.05	7.17	7.79	7.77	8.22	6.87	3.48	3.60	5.14	3.96	0.92	-4.59
INESTA	20.15	20.31	21.01	23.03	22.07	21.66	22.62	23.63	23.12	24.31	25.48	25.60	25.32	25.51	24.85
DEBTA	23.92	24.10	24.92	27.34	25.76	25.06	26.18	27.04	26.17	27.22	28.34	28.80	28.31	28.37	27.72
INVEST/A	5.92	5.64	4.86	3.45	4.13	4.99	4.92	4.77	4.74	4.42	4.08	4.10	4.24	4.02	3.02
LOGPPE	8.19	8.22	8.25	8.26	8.25	8.26	8.26	8.26	8.24	8.24	8.25	8.26	8.27	8.28	8.39
MOM	1.26	0.23	-3.78	3.57	2.09	0.58	1.48	3.03	0.89	0.45	1.16	1.68	0.27	1.03	1.94

Table(2.2). The table presents summary statistics for the control variables of the Russell 3000 index component. We print the average of each indicator by equi-weighting the single components each year. LOGSIZE is the natural logarithm of the market capitalization; B/M is firm's book value divided by its yearly market cap; ROE is the ratio of firm's net yearly income divided by the value of its equity; LEVERAGE is the ratio of debt to book value of assets; capital expenditures INVEST/A, is the firm's yearly capital expenditures divided by the book value of its assets; LOGPPE, is the natural logarithm of the firm's property, plant, and equipment; MOM is the average of returns on stock, for the 12 months' up to and including month t 1

In Table 2.2 we report summary statistics on stock returns and several control variables used in our subsequent tests. The dependent variable, $r_{i,t,s}$, in our cross-sectional return regressions is the monthly return of an individual firm i in month t, headquartered in state s. We use the following control variables in our cross-sectional regressions: $LOGSIZE_{i,q}$, given by the natural logarithm of firm i's market capitalization (price times shares outstanding) at the end of each quarter q; $B/M_{i,t}$, which is firm i's book value divided by its yearly market cap; $ROE_{i,t}$, which is given by the ratio of firm i's net yearly income divided by the value of its equity; $LEVERAGE$, which is the ratio of debt to book value of assets; capital expenditures $INVEST/A$, measured as the firm's yearly capital expenditures divided by the book value of its assets; $LOGPPE$, which is given by the natural logarithm of the firm's

property, plant, and equipment at the end of year t; $MOM_{i,t}$, which in turn is given by the average of returns on stock i, for the 12 months up to and including month $t-1$.

To assess portfolio exposure to the classical three Fama-French factors, we download the factors from Kenneth French's data repository (French (2020)) for the U.S. equity market. These factors are related to the Russell data set we employ for the financial analysis. Table 2.3 provides summary statistics on the three factor[6] characteristics: market return minus risk-free rate (Mkt-RF), small minus big (SMB) and high minus low (HML).

We collect data covering September 1990 to December 2020 on electricity consumption and weather futures prices to validate the materiality of our metric. We obtain time-series data on energy demand for 50 U.S. states at the monthly frequency from the U.S. Energy Information Administration. The U.S. classification considers four end-use sectors: residential (homes and apartments), commercial (offices, malls, stores, schools, hospitals, hotels, warehouses, and public assembly), industrial (facilities and equipment used for manufacturing, agriculture, mining, and construction), and transport.

Weather futures contracts are traded on the CME, and a majority of weather contracts are based on temperature. Temperature-related contracts insure the buy-

[6]For a comprehensive description, refer to Fama and Kenneth (1993)

Table 2.3. Summary Statistics, Fama French 3 Factors

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Mkt-RF	0.28	0.83	0.12	-3.68	2.28	1.49	0.14	1.31	2.57	0.96	0.08	1.09	1.63	-0.49	2.13	2.07	2.69
SMB	-0.02	0.18	-0.64	0.53	0.67	1.02	-0.37	0.01	0.48	-0.56	-0.48	0.73	-0.41	-0.39	-0.36	0.51	1.84
HML	0.73	0.86	-1.41	0.17	-0.23	-0.30	-0.70	0.71	0.18	-0.13	-0.84	1.62	-0.96	-0.87	-0.68	-2.92	4.15
RF	0.25	0.39	0.38	0.13	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.02	0.07	0.15	0.18	0.04	0.00

Table(2.3). The table shows the yearly average for the three Fama-French Factors. Source: Fama French repository (French (2020)). From the top to the bottom, the first, Market minus Risk Free(Mkt-RF). is the market excess return over the risk free. Small Minus Big (SMB) is the riturn of the least capitalized over the most capitalized. High minus low (HML) represent the spread between value and growth stocks. Risk Free(RF) is the US risk free rate.

ers against either excessive heat or excessive cold during a specified period of time. The two main temperature instruments are Heating Degree Day (HDD) contracts and Cooling Degree Day (CDD) contracts. The buyer of an HDD contract receives payments for cold days, defined as days with average temperature below 65°F; conversely, the buyer of a CDD contract receives payments for hot days, defined as days with average temperature exceeding 65°F. Contracts are available for eight geographically distributed cities across the U.S. These contracts are written on the observed temperature at a specific weather station near the contract city for a specific period. We select the same cities considered in Diebold and Rudebusch (2022) and Schlenker and Taylor (2021): Atlanta, ATL; Chicago, ORD; Cincinnati, CVG; Dallas Fort Worth, DFW; Las Vegas, LAS; Minneapolis St Paul, MSP; New York, LGA; and Portland, PDX.[7] Daily futures prices (end of day) for HDD and CDD contracts were obtained from Bloomberg, covering 2005 to 2020.

2.2.3 News and attention index

Concerning news indices, we employ data from several sources. From Google Trends we retrieve the Google Search Volume Index (SVI) from 2004 for a particular topic at the U.S. country and sub-national level.[8] The retrieved monthly index, $G_{s,t}$, represents the intensity of the topic search, on a scale from 0 and 100, in a certain region, s, from 2004 to the present. For each state, 0 represents a month with no searches on the topic whereas 100 is the month with the most searches in its history. Two states may peak at the same time; however,because the scale only compares time periods within a state, two states may have the same index value at the same time, without having the same actual search volumes. We download the Google SVI for the 50 states based on "climate change" and "temperature" search terms. Figure (2.1) shows the difference between the indices averaged at the U.S. level. We note

[7]Table 2.1 indicates which of the cities in our larger city sample have temperature derivatives available.

[8]Google makes the Search Volume Index (SVI) of search terms public via the product Google Trends (www.google.com/trends). Weekly SVI for a search term is the number of searches for that term scaled by its time-series average.

that searches for "temperature" display a clear seasonal pattern, which we de-trend using the methodology outlined in Choi et al. (2020).

Figure 2.1. Google Search Volume Index (SVI) - U.S. average

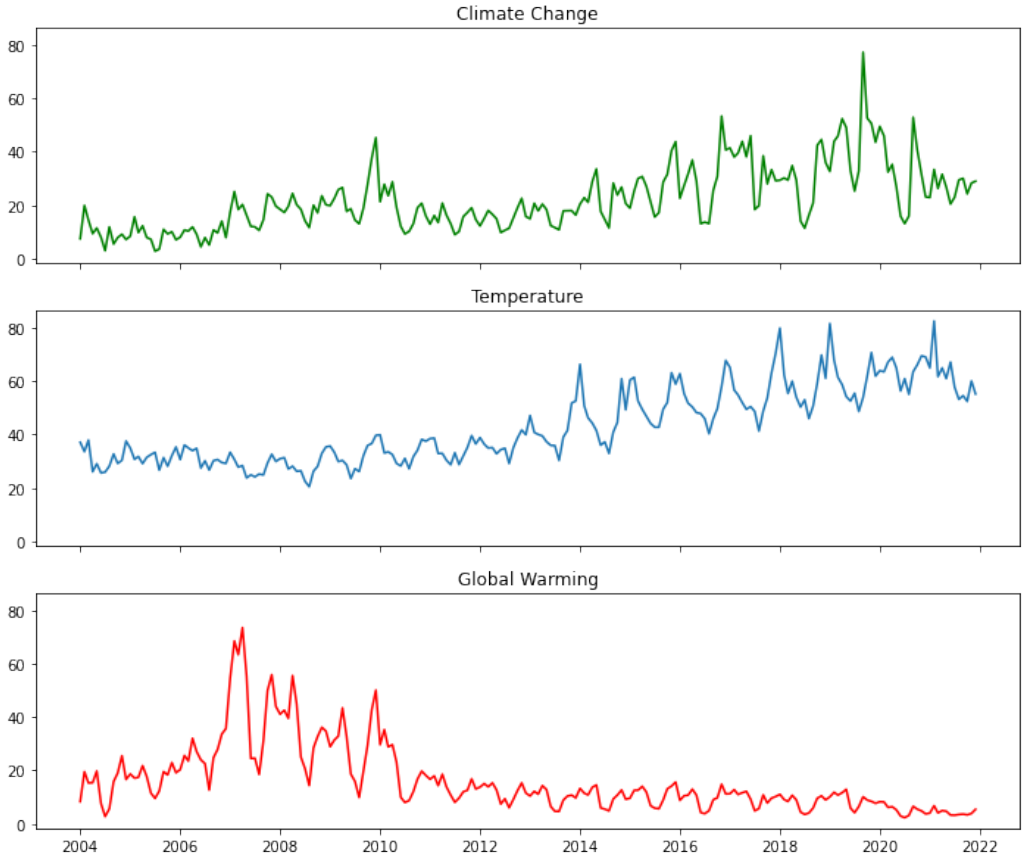


Figure 2.1 shows the monthly U.S. average of 3 Google Trends indexes, respectively "Climate Change", "temperature" and "Global Warming". Each month, the U.S. index is computed as $US_m = \frac{1}{50}\sum_i^{50} G_{i,m}$. Given that in each state the index assume values between 0 and 100, the US index is the simple average of the relative search in each month. This is just for illustration purpose given that in the analysis we employ state specific indexes.

We also use the WSJ news index created by Engle et al. (2020), which captures both physical and transition risks. The series they build is based on the assumption that any news about climate change is bad news. The news index is broken down by month, covering July 2008 through June 2017. When using this index, our sample is truncated to reflect this shortened time period.

2.2.4 Other data

In order to aggregate the state temperature index at the country level, we employ a population and gross domestic product (GDP) weighting method. We obtain the population and GDP series from the Federal Reserve Economic Database released by the Federal Reserve Bank of St. Louis. Population is available at the state level through the code 'POP', with a starting date of 1950. To match the monthly temperature and financial data sets, we interpolate the yearly frequency of the population data to obtain monthly or daily series. GDP is available through the code 'RQGSP', which returns quarterly real gross product for each state. We perform a similar interpolation method to obtain a monthly or daily series.

We examine firm-level exposure to temperatures by adopting data developed by Sautner et al. (2020). Their physical exposure measure consists of the proportion of time an earnings conference call centers around physical climate shocks. The data is available at a quarterly frequency and runs from 2000 to 2020. When using this measure, we average our temperature and attention data to the quarterly frequency.

2.3 On the evolution of excess temperature dynamics

Our aim is to represent the spatio-temporal variation in the distribution of temperature anomalies and, in particular, extreme temperatures, in a way which is more salient for financial markets. To do so, we expand upon prior temperature statistics to incorporate a more complete representation of abnormal temperatures by gradually building a statistic of temperature variability, named $TD\text{-}VAR_{,}$. With concrete and theoretical examples, we illustrate *why* our statistic is comprehensive in quantifying the extremes in the distribution of temperature. We end the section by describing our aggregation methodology used for our empirical analysis.

2.3.1 TD–VAR derivation

In the spirit of Donadelli et al. (2019) and Kotz et al. (2021), for a given location, we compute the variability in temperature deviation and subtract this value from its historical mean. This metric effectively captures changes in temperature extremes by means of shifts in temperature variability across time. It has been shown that a high degree of variability in temperature anomalies tends to be associated with more frequent extreme temperature events (Hansen et al. (2012)). Considering that the literature confirms that abnormal temperature anomalies cause economic and financial disruptions, we argue that deviations in temperature variability are a primary driver of market reactions, and this statistic improves upon alternative statistics used in the recent literature, such as changes in mean temperature and abnormal or extreme temperatures (Addoum et al. (2020), Addoum et al. (2021), Pankratz and Schiller (2021)).

To that end, we construct a set of location-specific temperature statistics. First, similar to Kotz et al. (2021), we define the daily change in mean temperature as:

$$TD_{s,d} = (T_{s,d} - \overline{T}_{s,d}), \tag{2.1}$$

where $T_{s,d}$ indicates the *maximum* temperature observed in location s on a certain day d in the year; and $\overline{T}_{s,d}$ represents the historical average temperature in the same location s and on the same day d over the period 1960–2005. A smoothing window of 5 days is used to calculate the historical average daily temperature.

Then, we define the monthly *temperature anomaly*, $\widetilde{TD}_{s,m}$, as the average of the temperature anomalies relative to the daily mean temperature within a given month. Analytically, this corresponds to:

$$\widetilde{TD}_{s,m} = \frac{1}{D_m} \sum_{d=1}^{D_m} TD_{s,d}, \tag{2.2}$$

where D_m is the number of days within month m.

Next, we calculate the monthly standard deviation in temperature anomalies in month m:

$$\sigma(TD_{s,m}) = \frac{1}{D_m}\sqrt{\sum_{d=1}^{D_m}(TD_{s,d} - \widetilde{TD}_{s,m})^2}. \tag{2.3}$$

Lastly, we calculate the deviation in temperature variability from its historical mean and obtain the *deviation in temperature variability:*

$$TD\text{-}VAR_{s,m} = \sigma(TD_{s,m}) - \overline{\sigma}(TD_{s,m}), \tag{2.4}$$

where $\overline{\sigma}(TD_{s,m})$ represents the historical average temperature variability recorded in location s at month m.[9] This second term distinguishes us from Donadelli et al. (2019), as their measure captures the dispersion of temperature variability against a historical level observed over the industrial revolution era, i.e. 1659–1759.

We emphasize that $TD\text{-}VAR$ can assume both negative and positive values. A negative value of $TD\text{-}VAR$ corresponds to a tighter distribution of temperature anomalies compared to historical realizations. Consequently, the likelihood of a large anomaly decreases. A positive value of $TD\text{-}VAR$ corresponds to wider variability of temperature anomalies and, by construction, an increase in the unconditional probability of large swings. A positive $TD\text{-}VAR$ can be the result of two distinct and non-concomitant patterns. First, where the temperature deviates strongly in one direction during elongated cold spells or heatwaves. Second, when day-to-day temperatures swing frequently between hotter– or colder-than-normal periods. In the next sections we describe where we observe these patterns in the data. Crucially, statistics that rely on shifts in mean temperature, $\widetilde{TD}$, and abnormal temperature are insufficient for describing changes in the unconditional probability of temperature.

[9]It is possible to compute an equivalent daily $TD\text{-}VAR$ over a rolling window l. First, we compute the standard deviation in temperature anomalies in location s at day d observed over the past l days: $\sigma(TD)_{s,d,l} = \frac{1}{l}\sqrt{\sum_{i=1}^{l} TD^2_{s,d-i}}$. Then, we compute the $\overline{\sigma}(TD_{s,d,l})$, the historical average temperature variability recorded in location s at day d with lag l. We then obtain the daily $TD\text{-}VAR$ defined as $TD\text{-}VAR_{s,d,l} = \sigma(TD_{s,d,l}) - \overline{\sigma}(TD_{s,d,l})$

2.3.2 Intuition on TD-VAR

Now that we have formally derived $TD\text{-}VAR$, we illustrate its ability to describe temperature extremes more comprehensively in comparison to using thresholds. Also, we discuss its advantages and various properties in the context of prior literature. Schlenker and Taylor (2021) and Addoum et al. (2020) use a form of TD by defining extreme temperatures as the number of days in a month where the temperature exceeds a given heat threshold. However, both heatwaves and cold spells can cause severe issues. For example, in early 2021, a winter storm hit Texas causing power cuts and, as lamented by the Financial Times, provoking disruptions to the global supply chain for chemical raw materials.[10] Addoum et al. (2021) correct this by accounting for both warmer and cooler temperatures, yet still use a form of a priori thresholding. When calculating temperature anomalies, thresholds are likely to vary substantially across time and location. $TD\text{-}VAR$ presents the most salient spatio-temporal information, providing a sufficient characterization of the distribution of temperature and concisely describing temperature extremes.

The primary benefit of using our metric versus other temperature measures such as TD, is its ability to capture temperature extremes without introducing idiosyncratic thresholds. Building on the literature that confirms the relationship between temperature anomalies and economic and financial disruptions, we claim that our measure $TD\text{-}VAR$ improves upon $\widetilde{TD}$ in representing noticeable characteristics of temperature anomalies and their economic and financial impacts. We show, through a set of exercises, the relative importance of $TD\text{-}VAR$ over $\widetilde{TD}$ in representing extremes in the temperature distribution.

We begin with a real-world example of occurrences of extreme temperatures and the average value of monthly temperature anomalies, $\widetilde{TD}$. Figure 2.2 represents the temperature conditions in Boston during the year 2020. The red bars represent the number of days that Boston experienced an 'extreme' temperature – defined as a

[10]The Financial Times, March 24th, 2021 - "Global supply chains face months of disruption from Texas storm"

temperature that exceeds 1.5 times the historical standard deviation. The blue line is the average monthly temperature anomaly, $\widetilde{TD}$. In January, the average temperature anomaly is substantial, but the frequency of extremes is low. The average temperature anomaly over December and April is modest but there is a high number of extreme days. In aggregate, the results show that a measure based on average anomalies does not adequately reflect the frequency with which a location experiences extreme temperatures.

We now turn to a theoretical example to illustrate why thresholds fail to represent temperature extremes. We consider a continuous response variable x characterized by a probability density function $\psi(x)$. The probability of experiencing an extreme value, X_{TD} is computed using the area above or below the thresholds k_{max} and k_{min}, which represent the values associated with the definition of an extreme.[11] Analytically, X_{TD} can be expressed as:

$$X_{TD} = \int_{k_{max}}^{\infty} \psi(x)dx + \int_{-\infty}^{k_{min}} \psi(x)dx, \tag{2.5}$$

where $\psi(x)$ is the probability density function of TD.

The two components we define, $\widetilde{TD}$ and $TD\text{-}VAR$, characterize shifts in the distribution parameters of TD. In the absence of climate change, $\widetilde{TD}$ is equal to 0, as is $TD\text{-}VAR$. Changes in $\widetilde{TD}$ represent the differences in the average realization of the temperature distribution; however, a positive $\widetilde{TD}$ does *not* imply an increase in the occurrence of extremes. In contrast, changes to $TD\text{-}VAR$ modify the entire shape of the distribution of TD, explicitly increasing the variability and the probability of extreme temperatures.
Simply put, changes in $\widetilde{TD}$ characterize shifts in the mean of TD while $TD\text{-}VAR$ accounts for a change in the variance of TD.

To continue our theoretical example, we present conceptual arbitrary thresholds typically used to represent temperature extremes. We assume that TD exhibits a

[11]This definition of temperature extremes is closely related to Pankratz and Schiller (2021), Addoum et al. (2020), and Addoum et al. (2021).

Figure 2.2. Extremes realization and TD average value

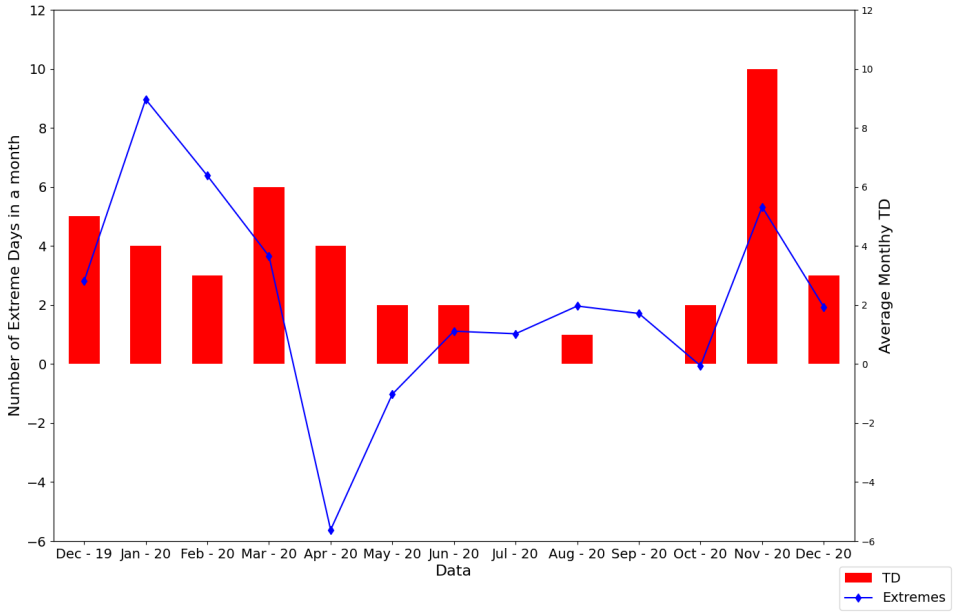


Figure (2.2): shows the historical evidence of extreme realizations and monthly average temperature deviation for the city of Boston in 2020. Red bars shows the number of day within a month the temperature experienced a an "extreme value", when the absolute value of temperature deviation higher than 2 standard deviations of historical distribution. The blue line shows the monthly average temperature deviation realization, Fahrenheit degrees.

normal distribution characterized by the parameters μ, which is set to zero, and σ^2 is set to the historical unconditional variance of TD.[12] Further, $\overline{\sigma}^2$ is fixed to 25 which is in line with the unconditional historical temperature data. Finally, we set k_{max} to 10°F and k_{min} to -10°F. In this baseline scenario, the amount of time a location experiences an "extreme" temperature is 4.6%, or 1 day in a month.

As there is no closed formula to formally calculate the amount of time spent above or below each threshold, we present a simulation that describes the interactions between the values of $\widetilde{TD}$ and $TD\text{-}VAR$.

Figure (2.3) shows the extreme probability, represented by the red shaded area for differing values of $TD\text{-}VAR$ and $\widetilde{TD}$. The central reference graph represents the historical distribution of temperature anomalies characterized by mean zero and the historical average volatility level. The historical average volatility level also implies that $TD\text{-}VAR$ is zero. Moving horizontally or vertically from the central plot represents nominal changes in $\widetilde{TD}$ and $TD\text{-}VAR$, respectively. The right column, where the temperature anomaly has a positive value, is associated with an increase in the occurrence of extremes *only* when $TD\text{-}VAR$ is non-negative. In contrast, positive values of $TD\text{-}VAR$ – in the first row – are always associated with a larger shaded area, even in the case of negative values of $\widetilde{TD}$. These instances illustrate the drawbacks of solely using $\widetilde{TD}$ rather than $TD\text{-}VAR$. In sum, using only the average of TD fails to consider potentially salient temperature realizations which $TD\text{-}VAR$ can better capture.

In the prior examples, we assumed $TD\text{-}VAR$ and $\widetilde{TD}$ to be independent when assessing $TD\text{-}VAR$'s relative importance from various theoretical perspectives. Here, we demonstrate how $TD\text{-}VAR$ correlates with other temperature measures. Table 2.4 shows the long-term relationship of $TD\text{-}VAR$ with other temperature factors, such as levels, deviation and variability. As already presented, there is no linear relationship between temperature anomalies and temperature levels. Deviation variability, however, shows a negative correlation coefficient of -0.8 due to the fact that

[12] μ equal to zero is a case where there is no global warming.

Figure 2.3. Effects on extreme, increase in TD and TDVAR

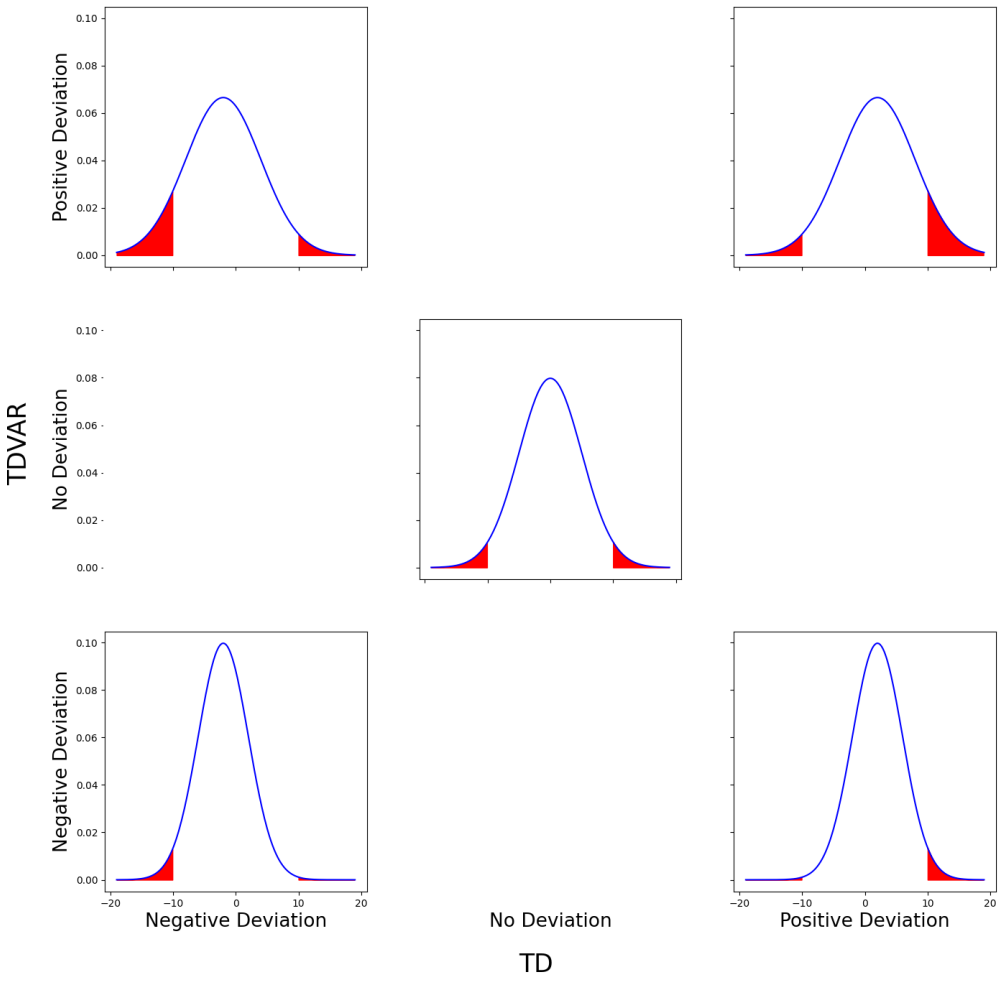


Figure (2.3): The table presents the interplay between increase in temperature deviation, TD, and deviation in variability, TDVAR. The columns entries present in the center, the historical level of TD that is equal to 0. The left-hand side column present a situation in which the average of $TD = +2$. The column on the left shows the situation in which the average of $TD = -2$. The entries on the rows show the comparable situation for TDVAR. The central rows, "Historical", present the situation associated with the historical variability level, $TDVAR = 0$. The Positive and Negative rows present respectively $TDVAR = 1, TDVAR = -1$. Red shaded area represent the probability of extreme, with threshold set at +10 and -10 Fahrenheit degrees.

variability is higher in winter and lower in summer, a fact well documented in the literature. When we turn our attention to $TD\text{-}VAR$, we note the negligible positive relationship with these other variables. We highlight that $TD\text{-}VAR$ does not display a seasonal pattern as the variability, and the relationship with variability, is absent.

Fundamentally, our examples highlight that temperature extremes can be better characterized using deviations in temperature variability. A key aspect of this argument is that it obviates the need for thresholds. We extend our theoretical explanation of the advantage of $TD\text{-}VAR$ over TD with realistic examples in the following sections.

2.3.3 City-level evidence

In this section, we supplement the prior theoretical examples with real-world, city-level records of $TD\text{-}VAR$ and $\widetilde{TD}$. We compute each value for 13 U.S. cities to further demonstrate the key differences between the two measures, analogous to Diebold and Rudebusch (2022). We select periods of time in these cities to graphically represent the temperature and temporal dynamics of its derivations. Furthermore, we highlight important differences after aggregating temperature data to the state level.

We obtain daily maximum temperature data from NOAA as our basis for our various offshoot measures. This data, from 1950 onward, are recorded at the airport stations of the 13 cities.[13] From the daily maximum level, we extract temperature, T, temperature anomalies, TD, temperature variability, $\sigma(TD)$, and deviation in temperature variability, $TD - VAR$. The former two metrics are defined at a daily frequency while the latter two are computed on a 30-day rolling window. Figure 2.4 displays the plots of these four statistics for the city of Atlanta between 2016 and 2018. Unsurprisingly, temperature T is characterized by seasonal changes, as illustrated in the top left diagram.

The random variable TD, shown in the top right panel, is generated after adjusting

[13]These temperature stations are the same as those used to compute the value of the temperature derivative contracts in Section 2.4.

Figure 2.4. TDVAR derivation for Atlanta

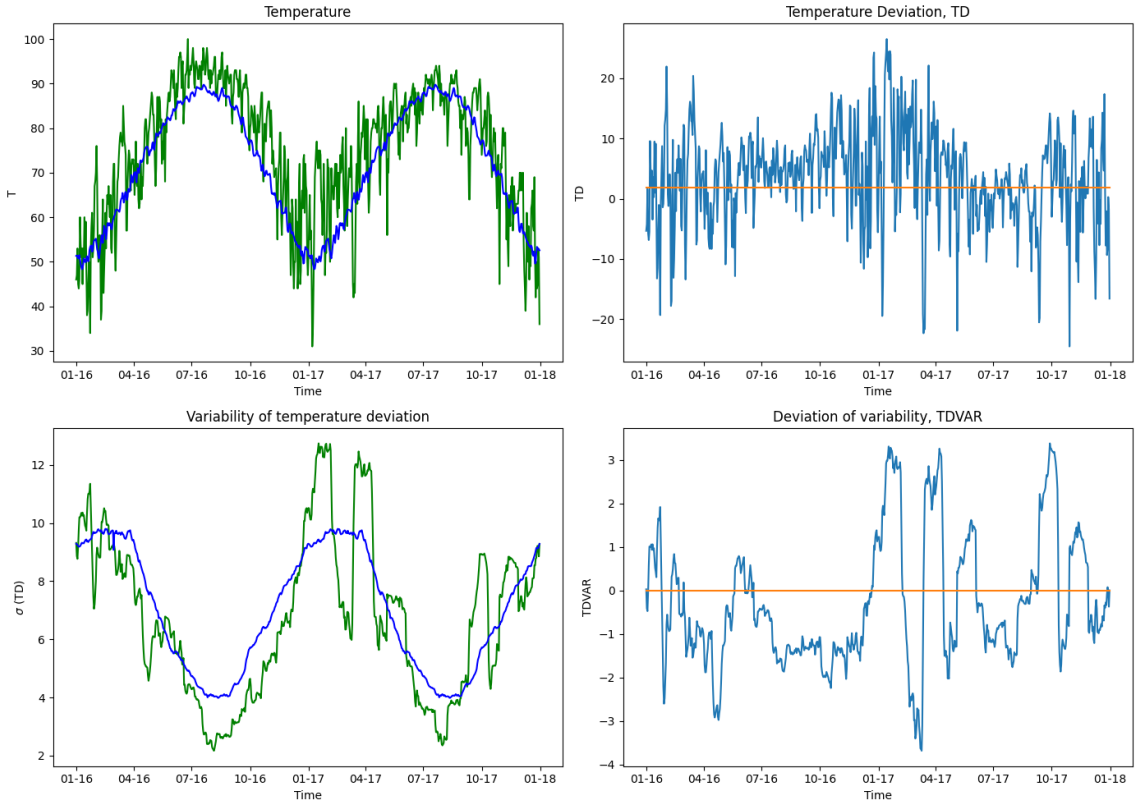


Figure (2.4) represents the steps embraced to extract from the temperature level the various component analyzed in the work. The top left figure shows in green the Temperature level, T, as well as the historical average of the Temperature level, $\overline{T}$. These components characterize the right hand side in equation (2.1). The top right panel displays daily Temperature Deviation in blue, TD, that is the left hand side of equation(2.1). The bottom left figure shows in green the volatility of temperature deviation, $\sigma(TD)$, defined by equation (2.3) and in blue its historical level, $\overline{\sigma(TD)}$. These two components represent the right-hand side of equation (2.4), whose outcome, $TD\text{-}VAR$, is shown in the bottom right figure, in blue.

temperature for seasonality.[14] This component exhibits a slow-moving long-term trend, shown in red, which represents the average warming due to carbon emissions. During the winter months, temperature anomalies are highly volatile compared to the summer months. The seasonality of TD volatility is made clear in the bottom left plot by calculating $\sigma(TD)$ using Equation 2.3. Lastly, deviations in temperature variability are derived in the bottom right diagram after removing the long-term seasonality/variability component. TD-VAR improves upon TD by removing the sluggish rise in temperatures and the seasonality component.

To further contextualize the plots in Figure 2.4, we explore the difference between TD-VAR and TD by highlighting one particular city location: Boston. We perform a rolling computation of $TD-VAR$ and TD at the daily level for a time window of 2015 through 2018, shown in Figure 2.5. The top and bottom panels of the figure contain the realizations of TD and TD-VAR, respectively. We exemplify a few moments in time to more realistically describe a link between TD and $TD-VAR$. There are three particular peaks in $TD - VAR$ for the city of Boston: June–July 2015, February–April 2016, March–April 2017. For the first peak, we note that TD is persistently high, portraying a heatwave, in comparison with the historical average – shown in red – during these months. The second peak of TD-VAR is related to large swings of temperature anomalies in both directions. The final high variability temperature event, beginning in March 2017, exhibits both persistently higher deviations and fluctuations around the historical average value. Combined, these instances show no direct link between the two measures at the daily level, which is corroborated by Table 2.4, as the correlation coefficient between $TD - VAR$ and TD is 0.12. To clarify the true disconnect between the two measures, we aggregate the values to the state level.

Figure 2.6 highlights the differences between the two temperature measures after deriving the monthly $TD - VAR$ and $\widetilde{TD}$ – Equation(2.2, 2.4) – for the state of

[14]Here, TD is comparable to measures in Pankratz and Schiller (2021), Addoum et al. (2020), and Schlenker and Taylor (2021). A caveat is that Pankratz and Schiller (2021) and Addoum et al. (2020) only use heatwaves.

Figure 2.5. TD and TDVAR characteristics

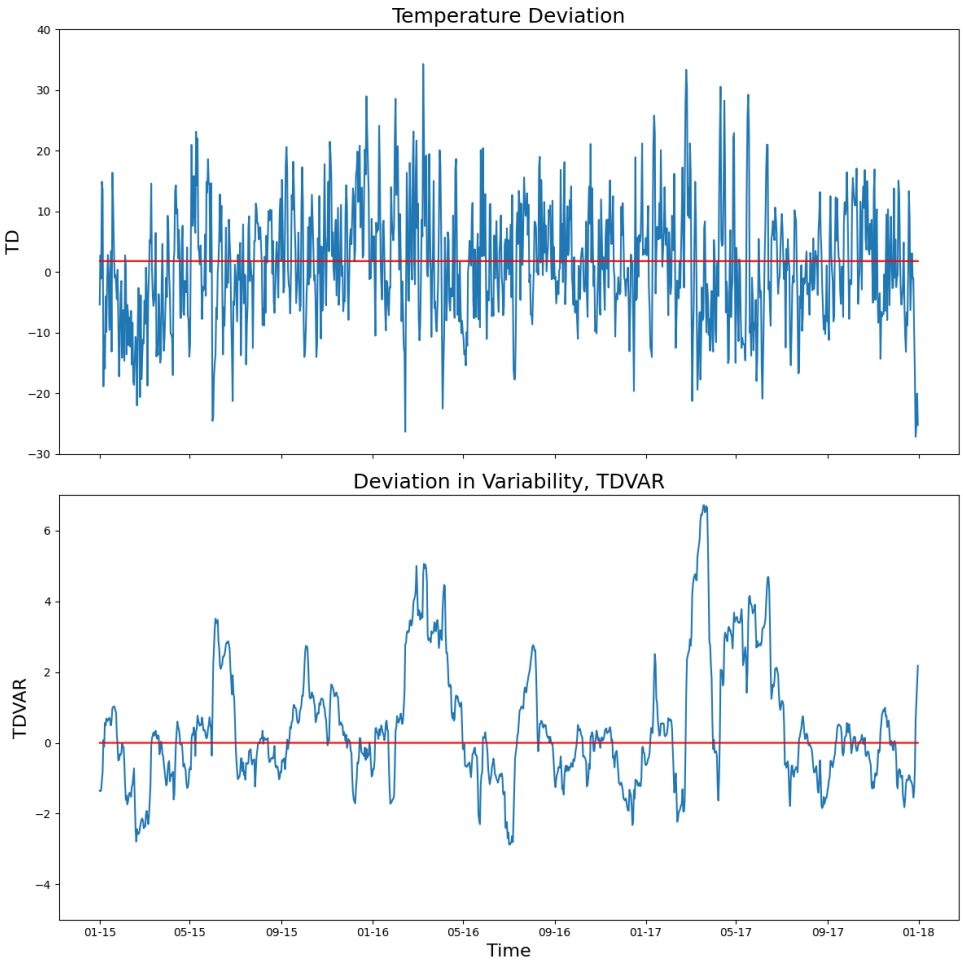


Figure (2.5) shows temperature characteristics for Boston in year 2015, 2016, 2017. The top figures shows the daily temperature deviation, TD, defined in equation (2.1). The red line presents the average TD over the period, that equals +1.8°F. The bottom figure present the daily TDVAR computed on a monthly rolling window following equation(2.4).

Texas. In the right panel of Figure 2.6, we plot monthly anomalies relative to historical temperatures, $\widetilde{TD}$, whereas the left panel illustrates the monthly deviation in temperature variability, $TD\text{-}VAR$. In May through June of 2018, there are high values of $TD-VAR$ in comparison to $\widetilde{TD}$, which is reversed in the months between October and December of 2019. This selected example illustrates that higher levels of TD are not tied to higher levels of $TD\text{-}VAR$, as the two metrics capture different aspects of the temperature distribution.

2.3.4 State and country aggregation

For our empirical analysis in the next sections, we use aggregated measures of $TD\text{-}VAR$ and $\widetilde{TD}$ to represent the spatio-temporal heterogeneity of temperature. Our underlying assumption is that firms and investors react to temperature shocks which are inherently local. We first aggregate gridded temperature anomalies to the state level because there is prior evidence that investors react to local weather shocks (Choi et al. (2020)). Additionally, due to data constraints and for reasons of simplicity, we assume that a firm's operational footprint is primarily located in the same state as its headquarters. To match country-level indices for our empirical analysis, we also combine the state-level aggregations.

We begin by computing state-aggregated temperature. We collect grid-level data from the BEST, which assigns a temperature field at a 1-degree resolution within U.S. land borders. The aggregation method can be formally defined by the temperature T of state s on day d as follows:

$$T_{s,d} = \sum_{i=1}^{N_s} w_i * T_{i,d}, \tag{2.6}$$

where $T_{s,d}$ is a weighted average of the temperature assigned to each grid cell $i = 1, \ldots, N_s$, and N_s is the number of parcel grids cells. We equally weight the cells by setting the weights w_i equal to $1/N_s$, which allows for a consistent measurement of average maximum temperatures at the state level. This methodology allows us

Figure 2.6. Monthly temperature characteristics, $\widetilde{TD}$ and $TD\text{-}VAR$

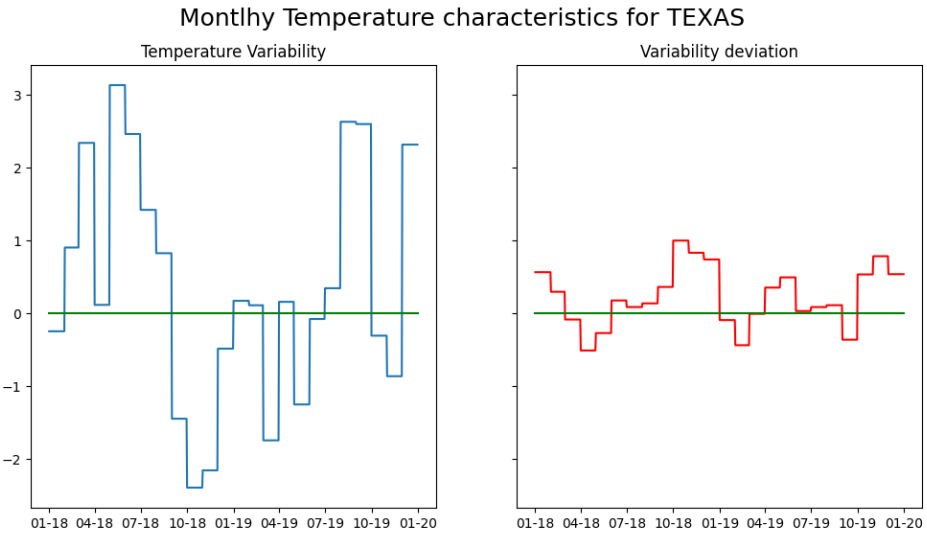


Figure (2.6): shows on the left panel monthly average temperature deviaton(TD 2.1) and on the right panel the monthly deviation in temperature variability ($TDVAR$, 2.4) for Texas between 2018 and 2020.

to derive $\widetilde{TD}$ and $TD\text{-}VAR$ of state s as described in expressions (2.1), (2.3), and (2.4). Our process of obtaining sub-national temperature data is similar to other papers such as Burke and Tanutama (2019). In Appendix A.1 we discuss alternative aggregation methods.

To derive U.S.-wide temperature factors, we aggregate each temperate volatility measure from its state counterpart. Aggregating data to the country level subsumes the rich sub-national heterogeneity of temperature data but allows us to match it to other country indices and data sources. We strictly define U.S.-wide $\widetilde{TD}$ and $TD\text{-}VAR$ as follows:

$$US_{\widetilde{TD},m} = \sum_{i=1}^{N_s} w_i TD_{i,m}, \tag{2.7}$$

$$US_{TD-VAR,m} = \sum_{i=1}^{N_s} w_i TD\text{-}VAR_{i,m}. \tag{2.8}$$

While there are multiple ways to aggregate temperature data, we consider each state

to be its own grid cell.[15] The measure for each state is then equally weighted by setting $w_i = 1/50$. In Appendix A.1 we discuss possible alternatives, such as weighting for the resident population or the GDP of a specific state.

2.4 Validation using electricity consumption and weather derivatives

Our prior examples reveal that the $TD\text{-}VAR$ measure compares favorably to TD in capturing the incidence of temperature extremes. Next, we test the salience and validity of our measure, $TD\text{-}VAR$, by investigating whether deviations in temperature variability are a relevant driver of energy consumption and prices in the weather derivatives market. This follows prior research by Campbell and Diebold (2005), who document that *unexpected* weather fluctuations can cause substantial pricing effects on the weather derivatives market and its players, such as energy producers and consumers. Given that $TD\text{-}VAR$ captures extreme fluctuations in temperature, we expect $TD\text{-}VAR$ to perform better than TD at accounting for variations in energy consumption and weather derivative prices.

We begin by examining the effect of $\widetilde{TD}$ and $TD\text{-}VAR$ for energy consumption. We obtain time-series data on energy demand at the monthly frequency from the U.S. Energy Information Administration for all states. Energy consumption is classified by

[15] Expressions (A.1) and (A.2) in Appendix A.1 describe alternative methods.

Table 2.4. Pearson Correlation Coefficient, Temperature drivers

	T	TD	$\sigma(TD)$	$TD\text{-}VAR$
T	1			
TD	0.32	1		
$\sigma(TD)$	-0.8	-0.11	1	
$TD\text{-}VAR$	-0.2	-0.12	0.17	1

Table (2.4). Sample period: 2005-2020. It is computed the Pearson correlation coefficient on the level of the variables. The entries represent Temperature level, temperature deviation (TD, 2.1), temperature variability (2.3), deviation in temperature variability (TDVAR, 2.4).

sector: residential, commercial, industry and transportation. Since energy consumption displays strong seasonal patterns, our analysis focuses on modeling short-run temperature shocks not captured by long-term trend analysis (Son and Kim (2017)). We link the observed seasonality of monthly demand (Bigerna (2018)) to the two components of temperature: anomalies and deviation in variability. We first run an ARMA (J,P) for each state s following Bigerna (2018):

$$Q_{s,t} = \sum_{j=1}^{J} a_j Q_{t-j} + \sum_{p=1}^{P} b_p \epsilon_{t-p} + \epsilon_{s,t} \tag{2.9}$$

where $Q_{s,t}$ represents the electricity consumption in state s at time t, J is the autoregression order and P is the moving average order. We then check the significance on the residual against $TD\text{-}VAR$ and TD respectively, and estimate a fixed effects model:

$$\epsilon_t = \beta_v * TD\text{-}VAR + \beta_t * \widetilde{TD} + \gamma_t + \eta_n + \epsilon. \tag{2.10}$$

Table (2.5) shows the resulting coefficients. We observe a positive and statistically significant β coefficient for the deviation in temperature variability, $TD\text{-}VAR$, in residential and industrial sectors and in the aggregate. A positive coefficient implies that, in a month characterized by high variability, the forecast value of electricity consumption exhibits a larger error relative to the best-fit value estimated through Equation (2.9). This error is inherently determined by the extent of the variability. The non-significant coefficient for $TD\text{-}VAR$ in the commercial sector suggests that the elasticity of electricity consumption is different for the residential and commercial sectors. This conclusion is supported by Zachariadis and Pashourtidou (2007) who find that the residential sector is highly reactive to weather conditions, as demand in the short term is inelastic to price. Taken together, our results confirm prior evidence of energy consumption being highly affected by weather conditions (Quayle and Diaz (1980)) and sensitive to large shifts in temperature variation. (Chang et al. (2016)).

Table 2.5. Estimation Results for energy consumption

	Residential	Commercial	Industrial	Total
$TD\text{-}VAR$	0.0054***	0.0006	0.0020**	0.0025***
	(0.0011)	(0.0006)	(0.0009)	(0.0005)
$\widetilde{TD}$	-0.0011	0.0013**	0.0004	0.0002
	(0.0008)	(0.0006)	(0.0004)	(0.0006)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
No. Observations	9000	9000	9000	9000
Cov. Est.	Clustered	Clustered	Clustered	Clustered
R-squared	0.0038	0.0034	0.0010	0.0027

Standard errors reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (2.5) The sample period is 2005-2020. The Table present the estimated coefficient for equation $\epsilon_t = \beta_v * TDVAR + \beta_t * TD + \gamma_t + \eta_n + \epsilon$ in the different sectors, Residential, Commercial, Industrial and the aggregated. Estimation is run trough a PanelOLS employing fixed effect for entities and time. Standard errors are clustered both at entity and time levels. The sample period is 2005-2020 for the 50 US state. $\widetilde{TD}$ and $TD\text{-}VAR$ are the state level temperature measure as defined in equation (A.2, A.1)

The relevant impact of weather on electricity demand has facilitated the creation of a market for weather derivatives. This market enables utility firms to hedge volumetric risk by trading the underlying risk driver – temperature – rather than the price of electricity (Jewson and Brix (2005)). We further validate our temperature measures by testing their association with city-level temperature derivatives prices.

We hypothesize that, if traders account for deviations in temperature volatility, $TD\text{-}VAR$ should capture more variation in weather derivatives prices than TD. To verify that our measure is relevant for weather derivative markets, akin to Diebold and Rudebusch (2022), we analyze futures contracts offered by the CME. The key benefit of this approach is that Schlenker and Taylor (2021) find that market participants accurately incorporate temperature anomalies through climate model projections. We extend this line of thought to confirm whether $TD\text{-}VAR$ is a driver of these contract prices. The first contract follows HDDs, which reflects the amount of heating required during cold days in winter. The second tracks CDDs that measure the necessary cooling required during hot days in summer. Therefore, CDDs have effective values

in summer and HDDs in winter. We strictly define CDDs and HDDs where T_0 is set at 65°F for a contract traded at the CME:

$$\begin{aligned} CDD_{i,m} &= \textstyle\sum_{d=1}^{D_m}(T_d - T_0)^+ \\ HDD_{i,m} &= \textstyle\sum_{d=1}^{D_m}(T_0 - T_d)^+. \end{aligned} \tag{2.11}$$

We use ordinal least square (OLS) regression analysis to investigate whether monthly average prices for CDDs and HDDs are affected by temperature – measured as temperature deviation and deviation in temperature variability. We estimate CDDs and HDDs separately with the following equations:[16]

$$\begin{aligned} CDD_{s,m} &= \beta_t T_m + \beta_e \widetilde{TD}_t + \beta_v TD\text{-}VAR + \beta_v \sigma(TD) + \gamma_m + \eta_s + \epsilon \\ HDD_{s,m} &= \alpha + \beta_t T_m + \beta_e \widetilde{TD}_t + \beta_v TD\text{-}VAR + \beta_v \sigma(TD) + \gamma_m + \eta_s + \epsilon, \end{aligned} \tag{2.12}$$

where T_m is the average daily temperature level minus 65°F degrees and $\widetilde{TD}$, $\sigma(TD)$, and $TD\text{-}VAR$ are defined in Section (2.3.1). For month and state fixed effects, we include γ_m and η_s, respectively. We only consider the constant term in winter given that the contract is not written on the maximum temperature of 65°F. We split the contract data into winter (October to March, inclusive) and summer months (April to September, inclusive).

[16] We use the derivatives defined in Section 2.3.3, and only consider the seven cities for which the derivatives are still traded.

Table 2.6. Estimation of Weather Derivates price driver

	CDD		HDD	
	(1)	(2)	(1)	(2)
T_m	22.262***	25.516***	-25.980***	-26.018***
	(1.7786)	(2.1067)	(0.8380)	(0.9349)
$TD\text{-}VAR$		4.0458**		3.5812***
		(1.9917)		(0.8282)
$\widetilde{TD}$		-11.082***		5.4309***
		(1.6592)		(0.6308)
$\sigma(TD)$		2.0248		19.595**
		(6.0450)		(9.2184)
α			326.87***	140.60*
			(11.508)	(79.420)
Estimator	PanelOLS	PanelOLS	PanelOLS	PanelOLS
No. Observations	438	438	542	542
Cov. Est.	Clustered	Clustered	Clustered	Clustered
R-squared	0.8807	0.9188	0.9501	0.9630

Standard errors reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (2.6). Sample period is 2015-2020. Estimation of model 2.12 for different specification. The dependent variabile is CDD and HDD respectively and the main independent variable is T_m that represent the maximum temperature minus 65°F, threshold level for futures contract traded at CME. Model (1) considers only T_m as regressor, that represent the underlying. (2) considers all the regresses.Estimation is run trough a PanelOLS employing fixed effect for entities and time. Standard errors are clustered both at entity and time levels

Table 2.6 shows the results for the two contracts using various temperature drivers. The first column of each panel includes the underlying temperature on which the contract is written, while the second column includes the other volatility measures. We show that T_m alone is able to explain 90% of monthly average price variance for CDDs in summer and 95% for HDDs in winter. An increase in temperature results in a decline in the price of CDDs, and vice versa for HDDs. Unsurprisingly, the magnitude of the coefficients is similar and of opposite sign, given that the derivative is dependent on the deviation from the 65°F threshold.

We then consider the remaining statistics: $TD\text{-}VAR$, $\widetilde{TD}$, and $\sigma(TD)$. We doc-

ument that historical variability, $\sigma(TD)$, has a large significant coefficient during the winter but no effect in summer. The result supports the findings in prior literature that temperature volatility is greater during these months.[17] The coefficients for TD-VAR are comparable across winter and summer, which is intuitive when recalling the option price effect of the volatility on the underlying asset. Higher deviations in temperature variability from the historical mean increase the probability of experiencing extreme temperatures and, consequently, increase the probability of exercising the option, thereby increasing the value of the weather derivative contract. This indicates that two cities with comparable average temperatures may face diverging weather derivatives prices when one city is characterized by higher temperature variability. Finally, we compare the coefficients of TD which have signs in the opposite direction to T_m. This suggests that traders assume temperatures will revert back to their historical levels when a city experiences higher temperature deviations. Collectively, we find that traders react negatively to increasing TD-VAR by establishing a higher price for the apparent risk, whereas an increase in TD implies a reversion to the mean for the market.

Our validation exercises strongly suggest that shifts in temperature variability, TD-VAR, are primary drivers of electricity consumption and the weather futures market. The results are consistent with Diebold and Rudebusch (2022) in demonstrating that refined measurements of temperature extremes can be consequential for financial asset prices. This confirmatory evidence also suggests that we are better able to characterize the reactions of market participants using deviations in temperature variability than by referring to temperature deviations alone. We continue this line of reasoning by asking whether this market response to TD-VAR has further implications for the stock market.

[17] Examining the seasonal component of temperature volatility, Campbell and Diebold (2005) and Benth and Benth (2007) document the higher values of temperature volatility during winter times.

2.5 Empirical analysis

2.5.1 Estimating temperature exposure

We analyze the effect of temperature deviation $\widetilde{TD}$ and deviations in temperature variability $TD\text{-}VAR$ on firm stock prices. Specifically, we are interested in examining whether the differential exposure of firms to deviations in temperature and temperature variability affect their stock price. Our expectation is that $TD\text{-}VAR$ will be associated with a collective reaction from investors as well as material shocks to a firm's performance, resulting in an aggregate decline in a firm's stock price. To empirically test this we use the Russell 3000, a broad index covering 3,000 major U.S. firms, and collect data on firm stock returns and headquarter locations. This framework benefits from the panel data characteristics of our sample, which is rich in both the cross-section and time-series dimensions. Moreover, it is possible to investigate the temperature dynamics at the industry level and include time– and firm fixed effects. We do not employ geography fixed effects since geographical differences are captured by the state-level temperature.

Table (2.2) in Section 2 provides summary statistics on stock returns and several control variables used in our analysis. The dependent variable, $r_{i,t,s}$, in our cross-sectional return regressions is the monthly return of an individual firm i in month t and headquartered in state s. We use the following control variables in our cross-sectional regressions: $LOGSIZE_{i,q}$, given by the natural logarithm of firm i's market capitalization (price times shares outstanding) at the end of each quarter q; $B/M_{i,t}$, which is firm i's book value divided by its yearly market cap; $ROE_{i,t}$, which is given by the ratio of firm i's net yearly income divided by the value of its equity; $LEVERAGE$, which is the ratio of debt to book value of assets; capital expenditures $INVEST/A$, measured as the firm's yearly capital expenditures divided by the book value of its assets; $LOGPPE$, which is given by the natural logarithm of the firm's property, plant, and equipment at the end of year t; $MOM_{i,t}$, which in turn is given by the average of returns on stock i, for the 12 months up to and including month

$t-1$. To allow for systematic differences in correlations across firms and over time, we include firms fixed effects η_t and year–month fixed effects ϕ_t. In this regard, our identification comes from states' variation in a given month.

In turn, we consider the effect of abnormal temperature, measured as the average daily temperature deviation within a month ($\widetilde{TD}$), and the effect of abnormal temperature variability, measured as the deviation in temperature variability from its historical mean in the same month ($TD\text{-}VAR$). This regression captures the impact of temperature on stock returns at the state level. Taking both abnormal temperature and abnormal temperature variability into account, we believe this measure provides a rough proxy for the climate change risk that a firm is exposed to at a given point in time.

Specifically, we estimate the following model:

$$r_{i,t,s} = \alpha + \beta_T * T_{t,s} + \beta_1 C_{i,t-1} + \phi_t + \eta_i + \epsilon_{i,t} \tag{2.13}$$

where $r_{i,t,s}$ measures the stock return of firm i in month t and headquartered in state s. T is a generic term that can stand in for either the deviation of daily temperature from its historical mean within a month ($\widetilde{TD}$), or the deviation of daily temperature variability from its historical mean in the same month ($TD\text{-}VAR$). The vector of firm-level controls C includes the firm-specific variables described earlier. We estimate these two cross-sectional regressions using panel OLS. In both model specifications, we cluster standard errors at the firm and year levels, which allows us to account for any serial correlation in the residuals and to capture the fact that some control variables are measured at an annual frequency.

We begin our analysis by asking whether temperature exposure affects the stock returns of Russell 3000 firms. Table 2.7 presents distinct estimates of the effect of abnormal temperature, Panel A, and abnormal temperature variability, Panel B. Specifically, once we control for firm and time period, as well as a battery of firm characteristics, the estimated effect of temperature deviation $\widetilde{TD}$ is economically

small and statistically insignificant when considering all firms (first column of Panel A). Nor do we find evidence of a significant relationship between an average firm's exposure to deviations in temperature variability, $TD\text{-}VAR$, and its stock returns (first column of Panel B).

Our empirical tests thus far indicate that exposure to temperature deviation and deviations in temperature variability has a minimal effect on stock returns for the *average* firm in the Russell 3000. To this point, our findings substantiate the results of Addoum et al. (2020), who document that temperature exposure is not an important driver of establishment-level sales growth. These findings are consistent with those of Dell et al. (2012), who show that the negative effects of temperature on aggregate economic growth are concentrated among developing countries, but tenuous in richer economies.

While there is some evidence that extreme temperatures have no effect (Addoum et al. (2020)), certain sectors of the economy may still exhibit sensitivity to abnormal temperatures. As such, we continue our analysis to examine whether firms in certain economic sectors are particularly sensitive. We use GICS codes to organize Russell 3000 firms into 10 sectors. In Table 2.7 Panel A, we rerun our yearly stock return regression for each sector. For all sectors except utilities (electric utilities; gas utilities; and multi-utilities), we continue to find economically and statistically insignificant estimates associated with exposure to temperature deviations. Months that are warmer or colder than expected are equally good and bad for all these industries. Utilities are a special case though, because they are tasked with providing enough energy over time as well as meeting instantaneous electricity demand, while juggling the costs associated with grid balancing and a continuous expansion of non-dispatchable renewable generation. As such, deviations in average daily temperature require utilities to invest more in emergency measures, such as increasing capacity and expanding demand–response investments to mitigate the effects of unexpected changes in daily temperatures. Accordingly, our analysis reflects that the effect of abnormal temperatures on utilities is economically important. The estimate indicates

that deviations of the daily temperature from the historical mean are associated with a 10.04 percentage-point decrease in utilities' stock returns, and that this effect is statistically significant.

In Panel B, we examine the effect of changes in the *distribution* of temperature by considering a deviation of daily temperature variability from its historical mean in a given month. As will become clear later, isolating the effect of changes in temperature distribution is decisive for understanding the temperature–stock relationship, and for qualifying some of the findings in previous studies that explicitly consider temperature extremes. Crucially, and in contrast to our estimates for the deviations in (average) temperature, we find that deviations in temperature variability significantly affect energy (oil, gas and consumable fuels; energy equipment), utilities, consumer staples (beverages, food products and tobacco; food and staples retailing; household and personal products) and consumer discretionary services (leisure products; textiles, apparel and luxury goods; hotels and restaurants; beverages; automobiles; and specialty retail).

Several channels may be at work to explain the negative impact of deviations in temperature variability on temperature-sensitive industries (Graff Zivin and Neidell (2014), Addoum et al. (2020), Addoum et al. (2021)).[18] Our findings are consistent with the consumer demand and labor productivity channels (Starr (2000); Graff Zivin and Neidell (2014)). Recall that $TD\text{-}VAR$ offers a general characterization of the unconditional probability of temperature extremes and, crucially, allows us to (i) simultaneously treat cold snaps and heatwaves as equally detrimental to economic activity, and (ii) capture day-to-day temperature swings between hot and cold. Using this measure, then, we find that many consumer-related sectors, including energy, are affected by changes in temperature variability. For example, large temperature swings can make shopping more or less difficult. Cold snaps and heatwaves can shift consumer demand patterns and may adversely impact what Starr (2000) calls "house-

[18]These papers examine various channels through which temperature affects economic output: manufacturing and labor productivity are sensitive to high temperatures, destruction of capital may occur at extreme temperatures, and consumer demand tends to drop, coupled with a decreased total labor supply.

holds' shopping productivity". Starr-McCluer provides empirical evidence consistent with these ideas using sector-level output data. This is also observable when considering macroeconomic output: Colacito et al. (2019) document that extreme heat in summer and autumn months affect U.S. GDP growth rates.

Table 2.7. Estimation for TD(A) and $TD\text{-}VAR$(B) on stock return

Panel A: Deviation of temperature

Dep. Variable: r	All	Ind	Energy	Health	IT	Utilities	Staple	C. Disc	Mat	Fin	Comm
$\widetilde{TD}$	0.0197	0.0510	0.1452	0.0852	-0.0076	-0.1004**	-0.0036	-0.0268	-0.0780	0.0119	-0.0101
	(0.9529)	(1.1508)	(0.9032)	(0.8592)	(-0.1381)	(-2.5604)	(-0.0500)	(-0.5041)	(-1.0290)	(0.3195)	(-0.1021)
LOGSIZE	-4.0221***	-4.3981***	-2.7935***	-4.0251***	-4.5562***	-3.1566***	-4.6041***	-5.1566***	-5.7560***	-2.8803***	-4.4655***
	(-27.758)	(-11.816)	(-5.6288)	(-8.2804)	(-9.5827)	(-7.0067)	(-8.3714)	(-14.519)	(-12.504)	(-11.977)	(-7.3882)
B/M	0.0028	0.0042	-0.0456*	-0.0224	0.0447*	0.0349	0.0355	0.0269	0.1293***	-0.0177*	0.0552**
	(0.4185)	(0.2098)	(-1.9141)	(-1.1148)	(1.7702)	(1.5437)	(1.5955)	(1.1453)	(4.6954)	(-1.7409)	(2.0166)
ROE	0.0421***	0.0360***	0.0181	0.0351***	0.0517***	0.0619***	0.0486***	0.0516***	0.0706***	0.0795***	0.0345***
	(12.679)	(4.1566)	(1.1995)	(4.4550)	(6.9696)	(2.8095)	(3.2003)	(8.2675)	(6.9395)	(6.4879)	(3.5929)
LEVERAGE	-6.228e-05	0.0014	-0.0255	-0.0001	0.0361***	-0.0033	0.0180	-0.0420***	0.0152	-0.0213**	0.0223
	(-0.1042)	(0.1429)	(-1.2480)	(-0.2092)	(2.6880)	(-0.1727)	(1.2436)	(-2.8260)	(0.9949)	(-2.2622)	(1.0244)
INVEST/A	0.0563**	0.0501*	0.0913	0.0618	0.0439	0.0126	0.1058*	-0.0246	0.0025	-0.1464*	-0.0403
	(1.9619)	(1.6797)	(1.5282)	(1.1325)	(0.6122)	(0.3603)	(1.9201)	(-0.8990)	(0.0481)	(-1.8887)	(-0.6033)
LOGPPE	0.3849***	0.9937***	1.7880*	-0.3065	-0.0303	0.1252**	0.9673***	0.8023***	0.7069*	0.4224**	0.9395**
	(3.7273)	(3.7276)	(1.8732)	(-0.8975)	(-0.0955)	(2.3325)	(3.7254)	(2.6222)	(1.7622)	(2.2345)	(2.4444)
MOM	-0.0431**	0.0235	-0.1375	-0.0134	-0.0499	-0.0545	0.0437	-0.0692*	0.0504	-0.2512***	-0.1031
	(-2.3345)	(0.5090)	(-1.2486)	(-0.2569)	(-1.1200)	(-0.5840)	(0.6600)	(-1.6582)	(0.8254)	(-5.3834)	(-1.2196)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	141827	26670	6731	17509	18058	6321	7441	18543	8911	22365	5380
Cov. Est.	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.0220	0.0237	0.0239	0.0138	0.0294	0.0210	0.0278	0.0355	0.0406	0.0279	0.0314

Panel B: Deviation of temperature variability

Dep. Variable: r	All	Ind	Energy	Health	IT	Utilities	Staple	C. Disc	Mat	Fin	Comm
TD-VAR	-0.0984	-0.2476	-0.9477**	0.4770	0.2354	0.3552**	-0.9432***	-0.6084***	0.0163	-0.0670	0.5953
	(0.0759)	(0.1541)	(0.4652)	(0.3478)	(0.2283)	(0.1538)	(0.2646)	(0.2236)	(0.2770)	(0.1357)	(0.4177)
LOGSIZE	-4.0223***	-4.3964***	-2.7972***	-4.0210***	-4.5571***	-3.1496***	-4.6158***	-5.1595***	-5.7561***	-2.8808***	-4.4402***
	(0.1449)	(0.3723)	(0.4961)	(0.4868)	(0.4755)	(0.4501)	(0.5489)	(0.3551)	(0.4605)	(0.2404)	(0.6057)
B/M	0.0029	0.0042	-0.0459*	-0.0232	0.0451*	0.0360	0.0353	0.0278	0.1289***	-0.0177*	0.0536*
	(0.0067)	(0.0200)	(0.0238)	(0.0201)	(0.0253)	(0.0226)	(0.0222)	(0.0235)	(0.0275)	(0.0102)	(0.0275)
ROE	0.0421***	0.0359***	0.0179	0.0350***	0.0518***	0.0616***	0.0486***	0.0516***	0.0705***	0.0794***	0.0347***
	(0.0033)	(0.0087)	(0.0150)	(0.0079)	(0.0074)	(0.0220)	(0.0152)	(0.0062)	(0.0102)	(0.0123)	(0.0096)
LEVERAGE	-6.133e-05	0.0012	-0.0261	-0.0001	0.0362***	-0.0043	0.0179	-0.0417***	0.0148	-0.0213**	0.0221
	(0.0006)	(0.0097)	(0.0204)	(0.0006)	(0.0134)	(0.0189)	(0.0145)	(0.0149)	(0.0153)	(0.0094)	(0.0218)
INVEST/A	0.0564**	0.0502*	0.0922	0.0575	0.0441	0.0134	0.1048*	-0.0245	0.0032	-0.1464*	-0.0430
	(0.0287)	(0.0298)	(0.0596)	(0.0547)	(0.0717)	(0.0351)	(0.0550)	(0.0273)	(0.0515)	(0.0776)	(0.0668)
LOGPPE	0.3850***	0.9949***	1.8210*	-0.3006	-0.0319	0.1170**	0.9734***	0.7981***	0.7175*	0.4214**	0.9523**
	(0.1033)	(0.2666)	(0.9533)	(0.3400)	(0.3172)	(0.0536)	(0.2603)	(0.3062)	(0.4015)	(0.1890)	(0.3845)
MOM	-0.0430**	0.0237	-0.1369	-0.0135	-0.0502	-0.0561	0.0468	-0.0698*	0.0516	-0.2513***	-0.1056
	(0.0185)	(0.0460)	(0.1104)	(0.0521)	(0.0446)	(0.0932)	(0.0662)	(0.0417)	(0.0611)	(0.0467)	(0.0843)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	141827	26670	6731	17509	18058	6321	7441	18543	8911	22365	5380
Cov. Est.	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.0220	0.0237	0.0243	0.0139	0.0295	0.0209	0.0295	0.0359	0.0405	0.0279	0.0317

Standard errors reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (2.7). The sample period is 2005-2020. All variables are defined in Tables (2.2) in the Data section. The independent variables include the deviation of daily temperature from its historical mean within a month (Panel A) or the deviation of daily temperature variability from its historical mean in the a month (Panel B). We use the Global Industry Classification Standard to identify a firm's sectoral affiliation. We consider the following sectors: Information Technology (IT), Health Care (Health), Financials (Fin), Consumer Discretionary (C. Disc), Communication Services (Comm), Industrials (Ind), Consumer Staples (Staple), Energy, Utilities, Real Estate (RE), and Materials (Mat). We refer to this document for an overview of the classification: http://www.msci.com/our-solutions/indexes/gics. We report the results of the panel regression with standard errors clustered at the firm and year levels. All regressions include month fixed effects and firm fixed effects.

The results in Panels A and B demonstrate the relevance of the temperature variability effect over and above the temperature deviation effect. This divergence

highlights the link between temperature and stock markets by documenting the actual impact of temperature variability on firms' operations. Deviation in temperature variability represents a more general depiction of temperature extremes than temperature deviation alone. Our measure is therefore a meaningful indicator of physical risk, which has material consequences for the stock price of firms.

In Table 2.8 we rerun the yearly stock return regression by splitting our sample into three time periods, illustrating the robustness of our findings across the following sub-periods: 2005–2009, 2010–2014, 2015–2020. We focus on a small group of sectors that display some interesting patterns: energy, consumer staples, and health care. The first two have significant exposure to $TD\text{-}VAR$ over the sample period 2005–2020, see Table 2.7 Panel B. Table 2.8 provides the estimates for $\widetilde{TD}$ and $TD\text{-}VAR$ for these three sectors. We report the results for all other control variables in the Appendix (A.1). Notably, the effect of $\widetilde{TD}$ remains insignificant in each sub-period, confirming the findings over the longer sample period in Table 2.7. Over time, the estimates of the effect of $TD\text{-}VAR$ on the energy sector decrease and then increase, and the estimates are virtually identical for consumer staples. There is no effect of $TD\text{-}VAR$ on the health care sector. These results confirm that exposure to temperature varies over time as the distribution of temperature and temperature variability changes over time (Lewis and King (2017), Alessandri and Mumtaz (2021)).

Table 2.8. Estimation for $\widetilde{TD}$ and $TD\text{-}VAR$, three sector, different periods

	Energy			Staple			Health		
Dep. Variable: r	2006-2010	2011-2015	2016-2020	2006-2010	2011-2015	2016-2020	2006-2010	2011-2015	2016-2020
$\widetilde{TD}$	0.267	0.1491	0.180	-0.324	-0.0739	0.3745	0.4770	0.1546	0.1036
	(0.4652)	(0.5246)	(0.7636)	(0.2646)	(0.2898)	(0.4061)	(0.3478)	(0.3990)	(0.5384)
$TD\text{-}VAR$	-0.4975*	-0.0863	-2.4626***	-1.0146***	-0.9042***	-0.8223***	0.4770	0.3278	0.5901
	(0.3652)	(0.5246)	(0.7636)	(0.2646)	(0.2898)	(0.4061)	(0.3478)	(0.3990)	(0.5384)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	6731	5171	3067	7441	5523	3252	17509	13700	9017
Cov. Est.	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.0243	0.0386	0.0621	0.0295	0.0260	0.0478	0.0139	0.0127	0.0161

Standard errors reported in parentheses
***1% significance,**5% significance, *10% significance,

Table(2.8). The sample period 2005-2020 is divided into three equal sub-periods. The independent variables include in turn the deviation of daily temperature from its historical mean within a month (first row) or the deviation of daily temperature variability from its historical mean in the a month (second row). We report the results of the panel regression with standard errors clustered at the firm and year levels. All regressions include month fixed effects and firm fixed effects

Given our initial sector-level evidence for the greater relevance of changes in temperature variability over temperature deviations, a natural follow-up question is whether financial market participants efficiently account for information on temperature deviation $\widetilde{TD}$ and deviations in temperature variability $TD\text{-}VAR$. To answer this question, we shift our focus to markets' *reactions* to both temperature deviation $\widetilde{TD}$ and deviations in temperature variability $TD\text{-}VAR$.

2.5.2 Reactions to local temperature information

In the previous section we perform an asset pricing factor analysis and examine the significance of the two temperature metrics. We continue our investigation into the relationships between these temperature metrics and expected stock returns by examining investors' reactions to state-level heterogeneity in temperature metrics. Specifically, we examine whether investors could reduce their exposure to temperature by focusing on local temperature information. To hedge against temperature, investors would buy (short-sell) stocks in states characterized by higher (lower) $\widetilde{TD}$ and $TD\text{-}VAR$, thus increasing (reducing) the prices of these stocks and reducing (increasing) their return.

Table 2.9. Quintile Transition Matrices: TD-VAR (A) TD (B)

Panel A

	Rebalanced Quintile				
	1	2	3	4	5
Quintile 1	52.29	24.82	14.10	5.54	3.25
Quintile 2	24.40	31.39	22.77	14.88	6.57
Quintile 3	12.41	24.28	28.98	22.95	11.39
Quintile 4	7.65	13.61	23.67	31.81	23.25
Quintile 5	3.25	5.90	10.48	24.82	55.54
Total	100	100	100	100	100

Panel B

	Rebalanced Quintile				
	1	2	3	4	5
Quintile 1	22.47	22.23	17.95	17.53	19.82
Quintile 2	20.48	20.78	21.08	19.34	18.31
Quintile 3	18.49	19.82	22.29	21.39	18.01
Quintile 4	19.52	20.36	20.72	20.90	18.49
Quintile 5	19.04	16.81	17.95	20.84	25.36
Total	100	100	100	100	100

Table(2.9). The sample period is 2006–2020. This table illustrates the frequency of states moving to another quintile of exposure. Panel A represents the transition matrix of the portfolio strategy when using $TD - VAR$ and Panel B represents TD. The left-most column of each panel is the beginning exposure quintile of the state. The other columns represents the exposure quintile of the state in the next month. Each number represents the percent of times a state moves from one quintile to another.

First, we sort states into quintiles based on their $\widetilde{TD}$ and $TD\text{-}VAR$ exposure, and sort stocks depending on companies' headquarter locations. Then, we form long–short spread portfolios: going long in the portfolio that includes states with the least exposure to $\widetilde{TD}$ and $TD\text{-}VAR$, and going short in the portfolio that includes states with the highest exposure to TD and $TD\text{-}VAR$. We examine whether the spread portfolios yield a statistically significant abnormal performance. If they do, this would suggest that investors do react to temperature information and, specifically, that they are pricing in the risk of temperature deviation, TD, and/or the risk of

deviations in temperature variability, $TD\text{-}VAR$,.

Table 2.10. Frequency of Rebalancing: TD-VAR (A) TD (B)

Panel A

Country	Trans.	Rank
FLORIDA	113	1
VIRGINIA	112	2
KENTUCKY	112	3
WISCONSIN	111	4
LOUISIANA	111	5
TENNESSEE	109	6
COLORADO	109	7
INDIANA	107	8
OHIO	107	9
NORTHCAROLINA	106	10
	⋮	⋮
ARKANSAS	94	40
IDAHO	94	41
MISSISSIPPI	92	42
OKLAHOMA	92	43
CALIFORNIA	91	44
KANSAS	91	45
MONTANA	87	46
OREGON	85	47
NORTHDAKOTA	82	48
SOUTHDAKOTA	80	49
ALASKA	79	50

Panel B

Country	Trans.	Rank
MAINE	142	1
DELAWARE	139	2
PENNSYLVANIA	138	3
OHIO	138	4
MISSISSIPPI	137	5
WISCONSIN	137	6
NEWHAMPSHIRE	137	7
ILLINOIS	136	8
NEWYORK	136	9
CALIFORNIA	135	10
	⋮	⋮
KANSAS	123	40
TENNESSEE	122	41
LOUISIANA	122	42
CONNECTICUT	121	43
TEXAS	121	44
NORTHDAKOTA	121	45
FLORIDA	121	46
SOUTHCAROLINA	119	47
WASHINGTON	118	48
WYOMING	118	49
IDAHO	115	50

Table (2.10). The sample period is 2006–2020. These tables show the frequency of states transitioning from one quintile to another for the portfolio strategy using $TD - VAR$. The first ten rows of the table show states that are often being rebalanced from one quintile to another while the second ten show states that remain in a quintile. The transition corresponds to the number of times a state rebalanced. The rank column denotes ranking of the number of times a state is rebalanced among the 50 states.

Our portfolio-sorting approach allows us to capture the heterogeneity in temperature across the U.S., similarly to Barber et al. (2001) and Hong et al. (2020). The trading strategy is constructed as follows. At the end of each month t, we rank states according to the specific realization of $\overline{TD}$ and $TD\text{-}VAR$ in month t. Separately, we

rank states based their $\widetilde{TD}$ and $TD\text{-}VAR$, and sort them into quintiles; we form $\widetilde{TD}$ and $TD\text{-}VAR$ quintiles separately. Each firm headquartered in a particular state is placed in one of the five quintiles for $\widetilde{TD}$ and then for $TD\text{-}VAR$. The first $\widetilde{TD}$ quintile portfolio, for example, consists of firms in those states with the lowest values for temperature deviation. We consider these firms to have the lowest exposure to temperature deviation. For each quintile, we compute the portfolio's post-ranking value-weighted monthly return. Next, we compute the long–short spread portfolio's monthly return. We repeat the process until we exhaust our sample period. This yields a time series of 167 spread portfolio monthly returns.

Table 2.11. Frequency of State in Quintile TD-VAR: 5th (A) 1st (B)

Panel A

	Freq.	Pct	Cum.
ALASKA	94	5.63	5.63
CALIFORNIA	74	4.43	10.06
OREGON	64	3.83	13.89
KANSAS	62	3.71	17.60
IDAHO	61	3.65	21.26
NEVADA	61	3.65	24.91
UTAH	56	3.35	28.26
MICHIGAN	50	2.99	31.26
TEXAS	49	2.93	34.19
ARIZONA	48	2.87	37.07
	⋮	⋮	⋮
FLORIDA	19	1.14	92.63
OHIO	18	1.08	93.71
NORTHDAKOTA	16	0.96	94.67
VIRGINIA	16	0.96	95.63
ALABAMA	15	0.90	96.53
MAINE	15	0.90	97.43
GEORGIA	12	0.72	98.14
SOUTHCAROLINA	12	0.72	98.86
HAWAII	11	0.66	99.52
MARYLAND	8	0.48	100
Total	1670	100	

Panel B

	Freq.	Pct	Cum.
NORTHDAKOTA	94	5.63	5.63
SOUTHDAKOTA	94	5.63	11.26
MONTANA	65	3.89	15.15
NEBRASKA	57	3.41	18.56
OKLAHOMA	52	3.11	21.68
MAINE	51	3.05	24.73
WASHINGTON	50	2.99	27.72
MINNESOTA	48	2.87	30.60
IOWA	47	2.81	33.41
ALABAMA	46	2.75	36.17
	⋮	⋮	⋮
RHODEISLAND	20	1.20	91.98
ARIZONA	19	1.14	93.11
CALIFORNIA	19	1.14	94.25
PENNSYLVANIA	19	1.14	95.39
NEWJERSEY	17	1.02	96.41
UTAH	16	0.96	97.37
VIRGINIA	16	0.96	98.32
ALASKA	10	0.60	98.92
WESTVIRGINIA	10	0.60	99.52
KENTUCKY	8	0.48	100
Total	1670	100	

Table (2.11). The sample period is 2006–2020. These tables show the frequency of states staying in a certain quintile. Panel A represents the 5th quintile when applying the portfolio strategy using $TD-VAR$ and Panel B denotes the same strategy using the 1st quintile. The first ten rows of each panel are the states that most consistently stay in either the 5th (A) or 1st quintile (B). The second ten rows of each panel are the states that least appear in either the 5th (A) or 1st quintile (B) portfolios. The frequency columns correspond to the number of times a state is found in a quintile portfolio. The percent column denotes percent of times the state is in the quintile portfolio out of all other states.

Table 2.12. Frequency of State in Quintile TD: 5th (A) 1st (B)

Panel A

	Freq.	Pct	Cum.
KANSAS	52	3.11	3.11
MONTANA	50	2.99	6.11
NORTHDAKOTA	49	2.93	9.04
NEVADA	48	2.87	11.92
IDAHO	47	2.81	14.73
OREGON	47	2.81	17.54
WYOMING	47	2.81	20.36
OKLAHOMA	46	2.75	23.11
CALIFORNIA	45	2.69	25.81
NEBRASKA	45	2.69	28.50
⋮	⋮	⋮	⋮
NORTHCAROLINA	23	1.38	90.12
RHODEISLAND	21	1.26	91.38
VIRGINIA	21	1.26	92.63
PENNSYLVANIA	20	1.20	93.83
DELAWARE	18	1.08	94.91
HAWAII	18	1.08	95.99
LOUISIANA	18	1.08	97.07
MARYLAND	18	1.08	98.14
CONNECTICUT	17	1.02	99.16
NEWJERSEY	14	0.84	100
Total	1670	100	

Panel B

	Freq.	Pct	Cum.
IDAHO	58	3.47	3.47
MONTANA	56	3.35	6.83
WASHINGTON	54	3.23	10.06
NORTHDAKOTA	53	3.17	13.23
NEBRASKA	51	3.05	16.29
OREGON	50	2.99	19.28
SOUTHDAKOTA	50	2.99	22.28
NEVADA	48	2.87	25.15
WYOMING	48	2.87	28.02
KANSAS	46	2.75	30.78
⋮	⋮	⋮	⋮
TENNESSEE	22	1.32	90.66
NEWJERSEY	21	1.26	91.92
DELAWARE	20	1.20	93.11
GEORGIA	19	1.14	94.25
CONNECTICUT	18	1.08	95.33
MARYLAND	18	1.08	96.41
LOUISIANA	17	1.02	97.43
RHODEISLAND	16	0.96	98.38
MISSISSIPPI	15	0.90	99.28
ALABAMA	12	0.72	100
Total	1670	100	

Table (2.12). The sample period is 2006–2020. These tables show the frequency of states staying in a certain quintile. Panel A represents the 5th quintile when applying the portfolio strategy using TD and Panel B denotes the same strategy using the 1st quintile. The first ten rows of each panel are the states that most consistently stay in either the 5th (A) or 1st quintile (B). The second ten rows of each panel are the states that least appear in either the 5th (A) or 1st quintile (B) portfolios. The frequency columns correspond to the number of times a state is found in a quintile portfolio. The percent column denotes percent of times the state is in the quintile portfolio out of all other states.

To fix ideas, at time t, the value-weighted return, R, of a quintile portfolio $p = \{1, 2, 3, 4, 5\}$ is:

$$R_{pt} = \sum_{i=1}^{n_{pt-1}} x_{it-1} r_{it}. \tag{2.14}$$

where r_{it} is the stock return of firm i-th at month t, and n_{pt-1} representing the

number of firms in the quintile portfolio p at month $t-1$. x_{it-1} represents the market capitalization of firm i divided by the total market capitalization of portfolio p at month $t-1$.

Table 2.13. Returns to Portfolios Sorted on $TD\text{-}VAR$ excluding four states

	Excess Return	**3-factor**	**4-factor**
Quintile 1	0.992***	0.275**	0.282**
	(2.913)	(2.275)	(2.379)
Quintiles 2–4	0.656**	0.001	0.001
	(2.117)	(0.023)	(0.013)
Quintile 5	0.56	-0.195	-0.190
	(1.545)	(-1.409)	(-1.382)
(1–5)	0.432	0.4703	0.395
	(x.xx)	(x.xx)	(x.xx)

t-stats reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (2.13). The sample period is from 2006 to 2020 These tables report the alpha (in percentage) to quintile portfolios sorted on $TD\text{-}VAR$. At the end of each month t, we sort states – excluding Alaska, California, North Dakota, and South Dakota – into quintile portfolios based on their $TD\text{-}VAR$, separately, using data up to month t. Returns for each quintile portfolio is the value-weighted returns of the firms headquartered in each state. Quintile 1 are those U.S. states with the lowest values of temperature variability $TD\text{-}VAR$. Quintile 5 are those countries with the highest values of temperature variability $TD\text{-}VAR$. We group the middle three quintile portfolios together by equal-weighting their respective returns and denote it as "Quintiles 2–4". We report the mean excess returns, alphas based on CAPM, three-factor model, and four-factor model. "(1–5)" reports the return spread between the top and bottom quintiles.

Table 2.14 reports the mean excess returns net of the U.S. risk-free rate. We also report portfolio alphas adjusted using the Fama-French three-factor model (Fama and Kenneth (1993)), which controls for the market factor as well as size and book-to-market factors; and the Fama-French-Carhart (Carhart (1997)) four-factor model, which includes Carhart's momentum factor. The middle three quintiles are grouped together by equal-weighting their respective returns.

Table 2.14. Portfolio Returns based on sorting metrics

Panel A: Returns to Portfolios Sorted on TD

	Excess Return	**3-factor**	**4-factor**
Quintile 1	0.878**	0.168**	0.166
	(2.585)	(1.670)	(-1.65)
Quintiles 2–4	0.697	-0.004	-0.003
	(2.182)	(-0.061)	(-0.052)
Quintile 5	0.813**	0.115	0.117
	(2.404)	(1.023)	(-0.001)
(1–5)	0.065	0.053	0.049

Panel B: Returns to Portfolios Sorted on $TD\text{-}VAR$

	Excess Return	**3-factor**	**4-factor**
Quintile 1	1.013***	0.293**	0.299**
	(3.001)	(2.467)	(-2.55)
Quintiles 2–4	0.6812**	0.009	0.008
	(2.191)	(0.180)	(0.153)
Quintile 5	0.737**	-0.003	0.00
	(2.195)	(-0.652)	(-0.001)
(1–5)	0.272	0.296	0.299

t-stats reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (2.14). The sample period is from 2006 to 2020. It reports the alpha (in percentage) to quintile portfolios sorted on TD (Panel A) and $TD\text{-}VAR$ (Panel B). At the end of each month t, we sort states into quintile portfolios based on their TD and $TD\text{-}VAR$, separately, using data up to month t. Returns for each quintile portfolio is the value-weighted returns of the firms headquartered in each state. Quintile 1 are those U.S. states with the lowest values of temperature deviations TD (Panel A); and lowest value of deviation of temperature variability $TD\text{-}VAR$ (Panel B). Quintile 5 are those countries with the highest values of temperature deviations TD (Panel A); and lowest value of deviation of temperature variability $TD\text{-}VAR$ (Panel B). We group the middle three quintile portfolios together by equal-weighting their respective returns and denote it as "Quintiles 2–4". We report the mean excess returns, alphas based on CAPM, three-factor model, and four-factor model. "(1–5)" reports the return spread between the top and bottom quintiles.

In the first column of both panel A and B in Table 2.14, the mean excess returns are largest in the first quintile portfolio, decline for portfolios 2-4, and increase slightly in the 5th quintile. The mean excess return for states in the bottom quintile portfolio of $\widetilde{TD}$ and $TD\text{-}VAR$ is 0.81% per month and 0.74% per month, respectively; for states in the top quintile, the mean excess return is 0.88% and 0.07%, respectively. The difference between quintile 1 (least exposed firms) and quintile 5 (most exposed firms)

is 0.07% per month in excess returns for $\widetilde{TD}$ and 0.28% per month in excess returns for $TD\text{-}VAR$. This is a respectable 47% return over the 14-year sample period. In columns (2) and (3), we report the portfolio alphas adjusted using the three-factor and the four-factor models, respectively. Notably, results are statistically significant only for $TD\text{-}VAR$ when including all four factors. These results suggest that markets hedge the risk from deviations in temperature variability $TD\text{-}VAR$, rather than the risk from temperature deviation $\widetilde{TD}$.

To the extent that local temperature anomalies might be concentrated in certain states, the significant return spread of our long–short portfolio could reflect compensation for exposure to a specific location or state risk. This is a plausible alternative as some states might be systematically exposed to continual increases or decreases in temperature deviation, or to deviations in temperature variability. Chronic exposure to abnormal temperatures would result in lower returns across the board for firms headquartered in these states. These firms could conceivably be less productive if the state in which they are headquartered is subjected to a barrage of temperature shocks. To explicitly control for location or state risk, we follow Barber et al. (2001), and construct two 5 X 5 transition matrices that illustrate the percent of times a state shifts to a different quintile of TD and $TD\text{-}VAR$, respectively, at month $t+1$. Table 2.9 reproduces the two 5 X 5 transition matrices. The starting quintile of the state is shown in the left-most column, while its quintiles at time $t+1$ are shown as the remaining columns. To shed more light on the influence of each temperature metric on the construction of the two portfolio strategies, we examine the frequency of states appearing in the first (least exposed) and fifth (most exposed) quintiles for both TD and $TD-VAR$ during our sample period. These figures are reported in Table 2.10.

Table 2.15. Returns to Portfolio Sorted on $TD\text{-}VAR$, sector specification

Panel A: Energy, Utilities, Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	1.114***	0.435***	0.439***
	(3.172)	(2.415)	(2.388)
Quintiles 2–4	0.7253***	0.162	0.16
	(2.833)	(1.473)	(1.446)
Quintile 5	0.678**	0.055	0.05
	(2.191)	(0.333)	(0.308)
(1–5)	0.436	0.38	0.389

Panel B: Energy, Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	1.314***	0.599**	0.679**
	(3.048)	(2.234)	(2.386)
Quintiles 2–4	0.734***	0.141	0.156
	(2.657)	(1.182)	(1.166)
Quintile 5	0.649*	-0.009	-0.19
	(1.9409)	(-0.048)	(-0.063)
(1–5)	0.665	0.608	0.869

Panel C: Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	1.083**	0.367	0.392
	(2.475)	(1.337)	(-1.456)
Quintiles 2–4	0.810***	0.229	0.227
	(2.916)	(1.646)	(1.622)
Quintile 5	0.766	0.100	0.106
	(2.104)	(0.443)	(-0.473)
(1–5)	0.317	0.267	0.286

t-stats reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (2.15). The sample period is from 2006 to 2020. It reports the alpha (in percentage) to quintile portfolios sorted on $TD\text{-}VAR$. At the end of each month t, we sort states into quintile portfolios based on their $TD\text{-}VAR$ using data up to month t. Returns for each quintile portfolio is the value-weighted returns of the firms headquartered in each state. We consider exclusively Energy, Utilities, Consumer Staples, Consumer Discretionary (Panel A); Energy, Consumer Staples, Consumer Discretionary (Panel B); Energy, Consumer Staples, Consumer Discretionary (Panel C). There are 57 energy companies headquartered in Texas, constituting approximately 45% of the Energy sector. These companies are excluded from the analysis. Quintile 1 are those U.S. states with the lowest values of temperature variability $TD\text{-}VAR$. Quintile 5 are those countries with the highest values of temperature variability $TD\text{-}VAR$. We group the middle three quintile portfolios together by equal-weighting their respective returns and denote it as "Quintiles 2–4". We report the mean excess returns, alphas based on CAPM, three-factor model, and four-factor model. "(1–5)" reports the return spread between the top and bottom quintiles.

The diagonal of the transition matrix in Panel B of Table 2.9 reveals that exposure to $TD - VAR$ is fairly persistent as some states do not move out of their initial quintile. The top left value in Panel B reports that a firm beginning in quintile 1 has a 52.29% chance of remaining in the first quintile. The values decline as we move into the middle quintiles, which suggests that exposure to $TD - VAR$ is persistent. Some states are either extremely exposed to deviations in temperature variability or not at all, and tend to stay near their quintile of exposure. Thus, exposure to $TD - VAR$ is dependent on the state. In contrast, exposure to temperature deviations (Panel A) is more erratic and less state-dependent. Fewer states are consistently exposed to extreme temperature deviations. For example, states in the fifth quintile are less likely to stay in this quintile, as the largest value in Panel A of Table 2.9 is 25.36% compared to its $TD - VAR$ counterpart of 55.54%. Results reported in Panel A of Table 2.9 therefore suggest that exposure to temperature deviation is less concentrated, or, equivalently, that most states are equally exposed to temperature deviations. This phenomenon is further illustrated in Table 2.10, which shows how often a state transitions from one quintile portfolio to another. Generally, changes in states' sorting order occurs 25%–40% more frequently when the ranking is based on their $\widetilde{TD}$ rather than their TD-VAR. The persistence of states' exposure to TD-VAR tells us that large changes in temperature variability are more concentrated vis-à-vis large changes in temperature. Temperature deviations occur uniformly across the U.S., with few states systematically subjected to large deviations in temperature. In contrast, a few states are continually unaffected by deviations in temperature variability.

These results highlights a significant difference between $TD - VAR$ and $\widetilde{TD}$. The idiosyncratic persistence of $TD - VAR$, particularly in the first and fifth quintiles, could influence the construction of our portfolio strategy. In fact, the tops and bottom of our $TD - VAR$ rankings are dominated by two states each: the Dakotas are consistently in the least exposed portfolio, Panel A in Table 2.11 for over half the sample period; Alaska and California are consistently in the most exposed portfolio,

Panel B in Table 2.11. Crucially, these four states are continually exposed to low and high values, respectively, for deviations in temperature variability. State rankings in the middle quintiles tend to vary more frequently across quintiles and are less bound to their original exposure. Table 2.12 displays the TD equivalent. One peculiarity is that North Dakota, Montana, Idaho, Kansas, Nevada and Wyoming occur in both the first and fifth quantiles of $\widetilde{TD}$. Crucially, this indicates large swings in temperature deviations in these states; this is aptly captured by our $TD\text{-}VAR$ measure, with some of these states appearing in the most exposed quintile in Panel B of Table 2.11. An issue with the infrequent rebalancing of states such as Alaska, California and the Dakotas is that the results of the portfolio strategy may be dominated by the firms in these states. This clustering in the first and last quintiles may drive the results and invalidate sorting based on temperature. To check the robustness of our results, we therefore focus our attention on $TD\text{-}VAR$ and perform the same portfolio strategy a second time, excluding Alaska, California, North Dakota and South Dakota. We believe that the removal of these four states is sufficient because the frequency of the remaining states in the first and fifth quintiles is comparable to their counterparts in 2.12. Results are reported in Table 2.13. When we calculate the total market-adjusted return over the 14-year sample period, the strategy produces a 66% net return – an increase of 16% compared to the previous long–short strategy. This result is driven by the smaller negative coefficient of -0.1094 in the fifth quintile. In contrast, the returns to portfolios sorted on temperature deviations $\widetilde{TD}$ are insignificant.[19]

To further explore investors' reactions to state-level heterogeneity in temperature metrics, we examine whether long–short portfolio strategies that only consider the temperature-sensitive sectors identified in the previous section exhibit monthly returns that differ from the all-industries average. We form long–short spread portfolios considering energy, utilities, consumer staples and consumer discretionary sectors (Panel A in Table 2.15). We then remove the utilities sector (Panel B) since electric and gas utilities tend to operate in the same state in which they are headquar-

[19]Results are not tabulated but are available from the authors upon request.

tered. For utilities, operational processes and consumer demand effects are relevant location-specific risks. By the same token, we remove the energy sector (Panel C). Furthermore, there are 57 energy companies headquartered in Texas, constituting approximately 45% of the energy sector. These companies are excluded from the analysis in order to reduce potential bias in those strategies where Texas appears in the first or last quintile. Panel A in Table 2.15 reports monthly returns of the long–short spread portfolios considering all four temperature-sensitive sectors. The strategy earns a remarkable 0.389% per month in excess returns adjusted for TD-VAR, and a negative 0.65% per month in excess returns adjusted for $\widetilde{TD}$ (Panel A2 in Table 2.16). We are, however, modest about our excess predictability results since our portfolios are poorly diversified across industries.

Table 2.16. Returns to Portfolio Sorted on TD, sector specification

Panel A2: Energy, Utilities, Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	0.881**	0.213	0.214
	(2.518)	(1.200)	(1.201)
Quintiles 2–4	0.755***	0.185	0.182
	(2.888)	(1.611)	(1.586)
Quintile 5	0.946**	0.293	0.298
	(2.782)	(1.569)	(1.601)
(1–5)	-0.058	-0.08	-0.084

Panel B2: Energy, Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	0.881***	0.241	0.239
	(2.711)	(1.508)	(1.499)
Quintiles 2–4	0.775***	0.178	0.177
	(2.769)	(1.467)	(1.452)
Quintile 5	0.923***	0.301*	0.304*
	(2.882)	(1.683)	(1.700)
(1–5)	0.665	-0.059	-0.065

Panel C2: Consumer Staples, Consumer Discretionary

	Excess Return	**3-factor**	**4-factor**
Quintile 1	0.946***	0.281	0.282
	(2.676)	(1.376)	(1.374)
Quintiles 2–4	0.822***	0.232	0.235
	(2.901)	(1.623)	(1.641)
Quintile 5	0.886**	0.226	0.227
	(2.411)	(1.067)	(1.101)
(1–5)	0.06	0.055	0.054

t-stats reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (2.16). The sample period is from 2006 to 2020. It reports the alpha (in percentage) to quintile portfolios sorted on $TD\text{-}VAR$. At the end of each month t, we sort states into quintile portfolios based on their $TD\text{-}VAR$ using data up to month t. Returns for each quintile portfolio is the value-weighted returns of the firms headquartered in each state. We consider exclusively Energy, Utilities, Consumer Staples, Consumer Discretionary (Panel A); Energy, Consumer Staples, Consumer Discretionary (Panel B); Energy, Consumer Staples, Consumer Discretionary (Panel C). There are 57 energy companies headquartered in Texas, constituting approximately 45% of the Energy sector. These companies are excluded from the analysis. Quintile 1 are those U.S. states with the lowest values of temperature variability $TD\text{-}VAR$. Quintile 5 are those countries with the highest values of temperature variability $TD\text{-}VAR$. We group the middle three quintile portfolios together by equal-weighting their respective returns and denote it as "Quintiles 2–4". We report the mean excess returns, alphas based on CAPM, three-factor model, and four-factor model. "(1–5)" reports the return spread between the top and bottom quintiles.

The returns from the long–short strategy increase substantially when removing the utilities sector, Panel B in Table 2.15. The monthly excess adjusted return moves to 0.869% for $TD\text{-}VAR$ and -0.065% for $\widetilde{TD}$ in Panel B2 of Table 2.16). When removing both energy and utilities, the long–short strategy for $TD - VAR$ generates an unremarkable adjusted excess return of 0.286% for $TD\text{-}VAR$, compared to the equivalent $\widetilde{TD}$ portfolio of 0.054% . Collectively, the large differential in the monthly gross excess returns between $\widetilde{TD}$ and $TD\text{-}VAR$ present strong evidence that investors hedge deviations in temperature variability rather than deviations in temperature. These results underscore a considerably larger reaction to deviations in temperature variability than deviations in temperature. Moreover, combined with the findings in Section 2.5.1, our results support the hypothesis that operating costs are a significant driver of state-specific temperature effects among energy and utilities sectors. For the consumer sectors, our results suggest that location-specific variability in temperature deviations has a lesser effect. The profitability effects among these industries tend to be driven by revenues rather than operating costs (Addoum et al. (2020)). Furthermore, the firms in these sectors have numerous brick-and-mortar stores spread across the U.S., which our aggregation method may fail to fully capture.

Overall, the results from our asset pricing tests show that $TD\text{-}VAR$ is a significant factor for firm stock prices when the shock occurs at a firm's headquarters, especially for the utilities and energy sectors. We expect our methodology to *underestimate* the true stock price reaction from temperature shocks due to data limitations. Using granular data on exact firm operations spatially overlaid with $TD\text{-}VAR$ would likely lead to even larger coefficients in our cross-sectional analysis, and greater adjusted returns using our trading strategy.

2.6 Identifying channels of price reaction

Our prior analysis strongly suggests that exposure to $TD\text{-}VAR$ has serious implications for firm stock prices and investors; however, we are agnostic as to the exact mechanism that dictates the price. Theoretically, there are two vectors at play. The first channel is investors' beliefs about companies that are exposed to temperature volatility. Heightened temperature variability acts as a "wake-up call" for investors, drawing attention to the risks of climate change, changing demand and simultaneously moving the equilibrium stock price of the exposed firm. The second, more direct channel, is the tangible, physical realization of the temperature shock on the firm's financial performance (Pankratz and Schiller (2021)).[20] The shock may lead to declines in revenues and profits, which are incorporated into the stock price.

2.6.1 U.S. country– and state-level attention

To test the first channel, we estimate the relationship between innovations in attention news indices and $TD\text{-}VAR$ and $\widetilde{TD}$. Innovations in attention are crucial, as expected news regarding climate change should not move the equilibrium prices of assets.[21] Investors should only react abnormally to unexpected news, thus we use the residuals from an autoregressive model with lag one for both country– and state-level indices. Investors should react to unexpected temperature swings by selling the exposed firm accordingly.

To test the relationship between unexpected changes in climate change attention and temperature volatility, we regress innovations indices on $TD\text{-}VAR$ and $\widetilde{TD}$ along with various fixed effects:

$$\epsilon_{AttentionIndex,s,t} = \alpha + \beta_T * TD\text{-}VAR_{s,t} + \beta_D * \widetilde{TD}_{s,t} + \rho_t + \gamma_s + \epsilon_{s,t}. \tag{2.15}$$

[20] For a deeper discussion on the exact channels of exposure, see Section 5.

[21] An investor's trading actions are 'conditioned' on their expectations of future climate shocks. However, unexpected shocks that are observed by investors may lead them to update their investments.

where the dependent variable is the AR(1) innovations in a particular country or state, s, index, ρ_t represents the time fixed effect and γ_s is the state fixed effects, when needed.

Aggregate attention is known to be a driver of asset prices. We adopt the innovations of the WSJ climate news index as our aggregate climate attention measure, as developed by Engle et al. (2020). Their research has sparked a growing literature on climate attention and the subsequent pricing of climate risks due to heightened attention. Engle et al. (2020) build the index from WSJ news articles that contain a discussion on climate change. Specifically, their measure captures the intersection between the news article text on climate change and the primary governmental or research source on which the article is based. Their assertion is that news articles on climate change are published more frequently when there climate concerns are apparent. Their narrative index connects an increase in news coverage of climate change with a heightened awareness of climate risks among investors. Direct investor attention toward a company, however, is difficult to capture and empiricists chiefly use indirect proxies (Da et al. (2011)). The WSJ index is an example of such an indirect proxy.

Table 2.17. Engle Index AR(1) Residuals

	1	2	3	4	5	6
$TD\text{-}VAR$	0.00140*		0.00134*	0.000568		0.000511
	(1.82)		(1.73)	(0.60)		(0.52)
$\widetilde{TD}$		-0.000154	-0.000134		-0.0000907	-0.0000801
		(-0.91)	(-0.75)		(-0.47)	(-0.40)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	No	No	Yes	Yes	Yes
No Observation	148	148	148	148	148	148
R-squared	0.11	0.09	0.12	0.17	0.17	0.17

t-stats reported in parentheses
***1% significance,**5% significance, *10% significance,

Table(2.17) The table shows the estimation for Equation 2.15 for the sample period 2005-2017. Here, the dependent variable for all columns is the innovations on the WSJ index. The estimation is conducted through panel OLS with fixed effects at the month and year level. Models (1) and (4) considers only $TD\text{-}VAR$ as the regressor, while models (2) and (5) use $\widetilde{(TD)}$. Models (3) and (4) consider both $TD\text{-}VAR$ and TD as regressors.

The results of the U.S. country-level regressions are found in Table 2.17. Here, we find evidence of investor reaction to $TD-VAR$ rather than to $\widetilde{TD}$. The coefficients of $TD\text{-}VAR$ in columns 1 and 3 are significant at the 10% level. When $TD\text{-}VAR$ increases by one standard deviation, there is an associated increase in $\epsilon_{AttentionIndex,s,t}$ by (0.2482 x 0.00140 =) 0.00035 or 58% of the mean of the innovations index. Adding both year and month fixed effects, however, cuts the coefficient by half, likely due to the variation captured by additional fixed effects. The coefficients on TD are non-significant and therefore indistinguishable from zero. The results indicate that $TD\text{-}VAR$ has a moderate relationship with the publication of climate-related news across the U.S. This association suggests that considerable unexpected temperature swings act as "wake-up calls" and grab the attention of news agencies and, subsequently, investors across the nation. The weaker relationship between $TD\text{-}VAR$ and the WSJ index is likely due to the spatial aspect of the shock. Realistically, shocks to $TD\text{-}VAR$ do not occur simultaneously throughout the U.S. but are instead highly localized. These attention-grabbing events will theoretically have a more substantial impact at the state level, which is represented in our state-level analysis.

Our results suggest that the number of unplanned articles written about climate change are positively and contemporaneously correlated with $TD\text{-}VAR$. However, we are agnostic as to the type of investor, as well as to whether investors do in fact read the articles that are published, as attention is a scarce resource. Nevertheless, the results thus far imply that investors are affected by elevated news coverage of climate risks, which lead to the pricing effects in Section 5.

Table 2.18. State specific Google SVI AR(1) residual

	Climate Change: Panel (A)			Temperature: Panel (B)		
	1	2	3	1	2	3
$TD\text{-}VAR$	0.77***		0.76***	0.73***		0.76***
	(0.24)		(0.25)	(0.17)		(0.16)
$\widetilde{TD}$		-0.05	-0.04		0.12**	0.13***
		(0.05)	(0.05)		(0.05)	(0.05)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	8850	8850	8850	8850	8850	8850
Cov. Est	Clustered	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.01	0.001	0.01	0.02	0.01	0.04

Standard errors reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (2.18). The sample period 2005-2020. Shows the estimation for Equation 2.15, on the sample period 2005-2020. SVI index is considered for the quotes "Climate change", Panel (A), and "Temperature", Panel(B). Estimation is run trough Panel OLS with fixed effects for entities. Standard errors clustered at the entity levels. Model (1) considers just $TD\text{-}VAR$ as regressor, model (2) employs $\widetilde{(TD)}$ and model (3) is run considering both $TD\text{-}VAR$ and TD.

Continuing our investigation of the first channel, we expect that less granular measures of investor attention, such as Engle et al. (2020), may not capture the rich geographic heterogeneity of $TD\text{-}VAR$ and $\widetilde{TD}$, as climate shocks are inherently regional (Alekseev et al. (2021)). Choi et al. (2020), for example, find that deviations in temperature only shift investor attention when they are substantial. To better capture the spatio-temporal distribution of temperature volatility, we associate temperature volatility shocks with local attention indices. We gather state-level Google SVI data on "Climate Change" and "Temperature", which should encompass reactions from retail investors and, to a lesser extent, from institutional investors (Da et al. (2011)). Specifically, we regress unexpected state-level SVI for each topic on the temperature volatility measures; the result is shown in Table (2.18).

The results of the state-level regressions show a positive and strongly significant relationship between $TD\text{-}VAR$ and index innovations. The coefficients for $TD\text{-}VAR$ are always significant at the 1% level and hover around 0.76 for both risk topics. TD is insignificant in the first panel, suggesting that deviations in temperature alone do not

raise attention. The TD coefficient, however, becomes highly significant in the second panel which is likely due to the higher relevance of the temperature topic to TD. A one standard deviation increase in $TD\text{-}VAR$ and $\widetilde{TD}$ leads to an increase in innovations by 35 and 25, respectively. The larger effect size of $TD\text{-}VAR$ substantiates our claim that this is the more salient metric. The results largely support our rationale that high intensity deviations from expected weather events affect investor attention.

We present U.S. country– and state-level evidence that temperature variability shocks, indicated by $TD\text{-}VAR$, are linked to investor attention. Our results are consistent with the interpretation that elevated $TD\text{-}VAR$ should "wake up" local investors and prompt an interest in climate risks, which then leads to a greater volume of searches in an exposed state. In comparison to the WSJ index, SVI is a revealed measure of attention as it implies that investors are directly paying attention to the topic (Da et al. (2011). The strong significance of the coefficients also strengthens the argument that attention is spatially dependent. In aggregate, the attention results show that $TD\text{-}VAR$ is a driver of unexpected attention toward climate change. This attention channel is one explanation of why investors reallocate their portfolios.

Table 2.19. Firm-Level Exposure to Temperature Shocks

	(1) All Industires	(2) Ex Util/Energy	(3) Util/Energy	(4) Cons Disc/Staples
$TD\text{-}VAR$	0.038***	0.032**	0.102**	0.020
	(0.014)	(0.015)	(0.050)	(0.019)
$\widetilde{TD}$	-0.006*	-0.005	-0.014	-0.004
	(0.003)	(0.003)	(0.013)	(0.005)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	65341	60589	4752	12046
R-squared	0.000	0.000	0.001	0.000

Standard errors reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (2.19) shows the estimation for equation 2.17, on the sample period 2005-2019. The primary dependent variable is $\epsilon_{NetExposure,i,t}$ which is obtained from equation 2.16. All four columns include both $TD\text{-}VAR$ and $\widetilde{(TD)}$ as well as firm and sector fixed effects. Each column includes a different set of industries as the sample: (1) includes all industries, (2) excludes the utilities and energy sector, (3) only views the utilities and energy sectors, and (4) includes only the consumer discretionary and consumer staples sectors.

2.6.2 Firm-level impact beyond attention

We investigate the second vector by distilling the realized impact of temperature shocks from firm-level earnings gleaned from conference call transcripts. We adopt a physical climate exposure measure by Sautner et al. (2020) who use these earnings call transcripts to develop a time-varying measure of firm-level exposure to physical climate change risks.[22] Earnings calls can be considered a more reliable gauge of climate exposure than management statements, such as annual reports, because the discourse includes both management and analysts. We argue that there is considerable overlap between attention and climate discourse as both investors and management may raise the issue during periods of global attentiveness. The results of Sautner et al. (2020) confirm this view, finding a positive relationship between the WSJ index developed by Engle et al. (2020) and their physical climate measure. To disentangle the material

[22]We limit ourselves to physical climate exposure because temperature shocks are a realized form of a physical climate shock.

impact of the temperature shock from the effects of attention, we obtain residuals from a regression of expected and unexpected attention on physical climate exposure created by Sautner et al. (2020).[23] Formally, we define this as:

$$PhysCCExposure_{i,t} = \alpha + \beta_1 * AttentionIndex_t + \beta_2 * \epsilon_{AttentionIndex,t} + \gamma_i + \epsilon_{NetExposure,i,t}. \tag{2.16}$$

Here we obtain $\epsilon_{NetExposure,i,t}$ as the remaining residual, which we argue contains the concrete impact of physical climate change exposure $PhysCCExposure$, beyond attention, for firm i at time t. We then regress our temperature shocks on this residual value to identify whether the shock had a tangible impact on firm-level operations. We formally define this by including the s subscript, which represents the state in which the firm is headquartered:

$$\epsilon_{NetExposure,i,t} = \alpha + \beta_T * TD\text{-}VAR_{s,t} + \beta_D * \widetilde{TD}_{s,t} + \gamma_i + \epsilon_{i,s,t}. \tag{2.17}$$

The main variation occurs at the state-quarter level, as this is where firms are exposed to time-varying temperature shocks. Our assumption here, similar to Section 5.2, is that the operational footprint of the firm is located in the headquartered state. While some literature, such as Pankratz and Schiller (2021) and Addoum et al. (2020), has granular data on firm establishments, we believe this generalization is still useful in capturing the spatial effects of the shock.

The results of Equation 2.17 with various sector samples are shown in Table 2.19. The first column of the table shows a positive, significant association with $TD\text{-}VAR$ and NetExposure for all firms in this sample. We interpret these results as $TD\text{-}VAR$ being associated with an increase in the proportion of an earnings call that discusses physical climate change exposure. Here, the conversation is not associated with broad climate attention, but rather with the *realized* physical exposure of the

[23] We only use the WSJ news index as it is the only attention measure used in Sautner et al. (2020). We adapt the WSJ index by taking the average over a quarter and finding the series' AR(1) residuals.

firm. When we exclude the utilities and energy sectors in column 2, the coefficient decreases slightly but retains its significance. The third column, which includes both the utilities and energy sectors, also displays a large positive coefficient between $TD\text{-}VAR$ and NetExposure. The utilities and energy sectors' high degree of sensitivity to $TD\text{-}VAR$ is consistent with prior evidence Section 5. In column 4, there is a positive, insignificant relationship between temperature shocks for the consumer discretionary and consumer staples sectors. The results are coherent with our Section 5 results, which suggested that the state-headquarter level inadequately represents the physical risk profile of consumer sector firms.

2.7 Conclusion

Extreme temperatures have been found to modulate financial markets. Furthermore, climate scientists have found that the distribution of temperature anomalies is becoming broader with an asymmetric lengthening of its tails (Hansen et al. (2012)). Using these facts as our motivation, we derive a metric, $TD\text{-}VAR$ which represents the deviation of the unconditional volatility from its historical level. We confirm the saliency of the metric on financial markets by using its monthly and annual realizations. At all stages, we compare our statistic to a widely accepted form of extreme temperature realizations: temperature anomalies or TD.

Through a set of empirical exercises, we demonstrate that shifts in $TD\text{-}VAR$ are primary drivers of: (1) energy consumption, (2) weather futures, and (3) U.S. stock markets. When we execute a hedging strategy by incorporating differential firm exposure, we find substantial market-adjusted returns suggesting excess return predictability. Finally, we investigate the underlying mechanisms and show that the observed pricing effects occur due to a combination of investor attention and firm-level repercussions as a result of changes to $TD\text{-}VAR$ rather than TD.

Our results have considerable implications for the energy and utilities sectors which are sensitive to day-to-day temperature variability as well as heat and cold

waves. While we find a moderated effect on consumer sectors, we believe that a larger effect size would be found with the inclusion of more granular footprint data to better identify firm exposure. We leave this for future research.

Chapter 3

Exploiting a risk driver of corporate bond overperformance: a statistical learning approach

3.1 Introduction

Understanding and forecasting future bond performances is a task usually underestimated in favor of stock forecasting. The main limitation arises from the lack of robust historical data that causes popular equity algorithms [1] to be less effective(He et al. (2021)). In addition, the variable researchers project poses additional challenges. The usual approach is to consider either total or excess return(see eg. Israel et al. (2017), Campbell and Taksler (2003), Lin et al. (2011)) or a derivation from the bond log price that takes into account the discount factor(Bianchi et al. (2021)). These procedures are not able to seize the underlying market invariant (Meucci (2005)) that drives bond performances. In this work, we firstly unravel a risk driver for bond excess return, for which we provide a coherent definition with the asset allocation process. Then we combine different statistical learning algorithms on a heterogeneous set of predictors to address future over-performances to overcome the unbalanced

[1] see eg. French et al. (1987); Scheinkman and LeBaron (1989); Fama (1990); Titman et al. (2004)

panel data framework. We analyze a corporate bond index, Bank of America Merrill Lynch Euro Corporate Investment Grade (BofA ML Euro IG, EG00), that is widely employed as a benchmark in asset allocation. Our approach starts by considering two sets of predictors, one at the bond instrument level and the second at the issuer level. The two collections are analyzed in a multivariate setting from a statistical learning algorithm that optimizes the right shape to match the target variable.

Our first contribution is the development of a risk driver for bond excess returns, defined as the bond performance above its duration-matched benchmark. Bond performances are driven by many factors such as yield, risk-free rate, credit spread, and bond duration, which are usually considered as explanatory variables(Bianchi et al. (2021), Hong et al. (2012)). In the process of asset allocation within asset management companies, corporate bond selection constitutes a downstream process, that usually follows indications coming from the Strategic Asset Allocation and the Tactical Asset Allocation division. The focus then should not be to forecast bond total return, given that macro factors and short-term movements are already taken into consideration at a higher level and a specific indication for the sector or duration bucket in which to invest is given. It is then crucial to assess which bond will perform better within a certain duration bucket and for such a reason, we define the bond excess return as the total return net of the average total return of the bonds in the same duration bucket. To disentangle the risk driver for the excess return we build up the decentralized return attribution (Carr and Wu (2019)) and we dub the Idiosyncratic change in Asset Swap spread (ASW), a measure tightened to the change in firm credit riskiness. We demonstrate that Idiosyncratic change in ASW is a market invariant(Meucci (2005)) and that it correctly models excess returns.

Our second contribution regards the adoption of a multivariate time series framework, that takes into account different bond characteristics to evaluate future performances. Following works by He et al. (2021), Bianchi et al. (2021) and Lin et al. (2018) we analyze a multivariate framework to project the risk driver and then extrapolate bond excess returns from which we can compose portfolios. In this framework, we com-

pare the performances at the portfolio level as well as the capability of the statistical learning algorithm to address the entire return distribution.

Our third contribution regards the number of entities an algorithm can analyze. Large data sets present the drawback of missing values, that prevent the analysis of the entire universe. Adding an explanatory variable with missing values for certain instruments forces us to discard these instruments even if this is the only missing value. Starting from the simple mean combination approach(MC) proposed by Stock and Watson (1998), large attention has been paid to combining the forecast of different models up to an iterative procedure (Lin et al. (2018)). We take advantage of these procedures to overcome missing data, by aggregating just the available measure and thus be able to enlarge the share of issues forecasted.

The last analysis we employ regards climatic variables, considering the transition risk component. Transition risk variables concern the risk of a shift in consumer preferences or regulatory tightening that could affect polluting firms. Carbon emissions are the primary driver for transition risk, especially when considering their variation (Bolton and Kacperczyk (2020)). Analysis on the bond universe find lower return for polluting firm (Duan et al. (2020)) or a green bond premium (Zerbib (2017), Reboredo and Ugolini (2020)) when examining green bond. We start by considering a change in CO_2 Emission that is employed as the only factor for transition risk. From a physical side, just weather-related events have been partially linked to yield spread(Zhang and Zhu (2021)). Another source of climate risk steams from the climate news index literature (see eg. Engle et al. (2020), Faccini et al. (2021), Ardia et al. (2020)) with preliminary work concerning the bond market (Huynh and Xia (2021)) that however needs a time series approach for a proper application .

3.2 Exploiting Excess Return

The first objective of this work is the exploitation of a risk driver for bond returns overperformance, that we define as the return of a bond over its benchmark. The usual approach in literature is to consider the return over a risk free rate (Cochrane

and Piazzesi (2005); Ilmanen (1995)). We deal with the typical procedure within an asset management company that faces an investing problem with a defined exposure to a certain asset, in this case corporate bond, and thus has already settled a risk profile over the risk free rate. In addition, we allow to split the asset class universe in N_k duration buckets k, that are associated with different risk profiles. Each duration bucket comprises all the bonds B_i characterized by a Duration D_B that is between the duration bucket limits, minimum, min_k, and a maximum, max_k.
We derive the excess return, er_B, for a bond in a duration bucket k, by subtracting from the total return, $r_{B_i,k}$ the average total return of the bonds in the same duration bucket k. It is formally defined as:

$$er_{B,k} = r_{B,k} - \overline{r_k} \tag{3.1}$$

where $\overline{r_k} = \dfrac{\sum_j r_{B_j}}{J}$ is the average return of bonds in the same duration bucket.
In order to disentangle the driver for bond overperformance we start from the bond decentralized return attribution investigated by Carr and Wu (2019), that decompose the bond's total return.

We start the derivation recalling that considering a bond characterized by a predefined stream of payments Π quoted on the market with a price B_t, it is possible to derive the Yield To Maturity (y) by solving the equation:

$$B_t = \sum_{j=1}^{I} exp(-y_t \tau_t)\Pi_j \tag{3.2}$$

where Π_j is the of payment at time j.

We then consider the instantaneous percentage return, dB/B, that represents the change in bond market price, as a function of its own Yield To Maturity as,

$$\frac{dB_t}{B_t} = \frac{\partial B_t}{B_t \partial t}dt + \frac{\partial B_t}{B_t \partial y}dy + \frac{1}{2}\frac{\partial^2 B_t}{B_t \partial y^2}dy^2 + o(dt) \tag{3.3}$$

where $o(dt)$ represents sensitivities of higher order that are negligible in this context. Combining Equation (3.3) and (3.2), allows to derive analytically the three classical sensitivities of bond prices.

$$\frac{\partial B_t}{B_t \partial t} = y_t, \quad -\frac{\partial B_t}{B_t \partial y_t} = \sum_{j=1}^{N} w_j \tau_j = D, \quad \frac{\partial^2 B_t}{B_t \partial y^2} = \sum_{j=1}^{N} w_j \tau_j^2 = C \tag{3.4}$$

where w_j represent the weight associated with the cash flow at time j. [2]

As emphasized by Carr and Wu (2019), assuming diffusive yield movement, it's possible to derive the bond annualized instantaneous return as

$$\frac{dB_t}{B_t dt} = y_t - D\frac{dy}{dt} + \frac{1}{2}C\frac{dy^2}{dt} \tag{3.5}$$

and, considering the flow of time, also the composite return over the period of interest dt, expressed as fraction of a year.

$$\frac{dB_t}{B_t} = y_t dt - D * dy + \frac{1}{2}C * dy^2 \tag{3.6}$$

Finally, the expected value in the real-world probability space, P, it is possible to obtain

$$\mathbb{E}_t^{\mathbb{P}}\big[\frac{dB_t}{B_t dt}\big] = y_t - D * \Delta_t + \frac{1}{2}C * \sigma_t^2 \tag{3.7}$$

where the second component, $\Delta_t = \mathbb{E}_t^{\mathbb{P}}[dy_t/dt]$, represents the expected change on the yield to maturity, whereas $\sigma_t = \mathbb{E}_t^{\mathbb{P}}[(dy)^2/dt]$ is the conditional variance of the yield process. If the expected variation in the yield to maturity, dy, is null, implying no shift in the yield curve, the only gain the bondholder receives is linked to the carry component, y, standing for the stream of payment expected[3].

Now we move on from decentralized return attribution and consider change in yield

[2]The derivation of the sensitivities implies that the weights are the present value of the specific stream with respect to the bond present value: $w_j = \frac{exp(-y_t\tau_t)\Pi_t}{B_t}$

[3]In the case the bond is at par, it's equal to the coupon, otherwise it encompass also the put to the par effect

to maturity component, dy. The Yield to Maturity is composed by two part, a Risk Free component and a Credit Risk one. Trough the Asset Swap Spread(ASW) it is possible to decompose the Yield to Maturity and its variation as this two components,

$$\begin{aligned} y_t &= rf_t + ASW_t \\ dy_t &= drf_t + dASW_t \end{aligned} \tag{3.8}$$

and then, employing the linearity property of the expectation, we decompose the expected change in dy as

$$\mathbb{E}_t^{\mathbb{P}}[dy_t/dt] = \Delta rf_t + \Delta_{ASW,t} \tag{3.9}$$

that is possible to plug in Equation (3.7). Recalling that $r_{B_i,k}$ coincides with the definition in Equation (3.7), we merge Equation (3.1) and (3.7) to obtain

$$er_{B_i,K} = y_{i,t} - D_i\Delta_i + \frac{1}{2}C_i\sigma_i^2 - \frac{1}{J}\sum_{j=1}^{J}(y_{j,t} - D_j\Delta_j + \frac{1}{2}C_j\sigma_j^2) \tag{3.10}$$

In addition, from Equation (3.8) it is possible to substitute Δ with $\Delta_{rf} + \Delta_{ASW}$ [4] in Equation (3.10), and, by properly modelling the terms, we rewrite (3.10) as follow

$$\begin{aligned} er_{B_i,K} = (y_{i,t} - \frac{1}{J}\textstyle\sum_{j=1}^{J} y_{j,t}) + \\ -(D_i * \Delta_{rf_i} - \frac{1}{J}\textstyle\sum_{j=1}^{J} D_j * \Delta_{rf_j}) + \\ -(D_i\Delta_{ASW_i} - \frac{1}{J}\textstyle\sum_{j=1}^{J} D_j\Delta_{ASW_j}) + \\ +\frac{1}{2}(C_i\sigma_i^2 - \frac{1}{J}\textstyle\sum_{j=1}^{J} C_j\sigma_i^2) \end{aligned} \tag{3.11}$$

It is possible to extrapolate four main components from the previous equation:

- Excess Yield to Maturity: $ey_i = y_{i,t} - \frac{1}{J}\sum_{j=1}^{J} y_{j,t}$, is the excess return that the bond would gain if all others components don't change in the period of analysis. In this context, this is deterministic.
- Relative change in risk free: $D_i * \Delta_{rf_i} - \frac{1}{J}\sum_{j=1}^{J} D_j * \Delta_{rf_j}$ is the difference

[4] where rf is the risk free rate interpolated at the precise bond duration.

between the change in the risk free associated with the bond duration and the average change in the risk free component of the bonds in the same duration bucket. In this case we won't lose any important information by setting this element to 0 because all the bonds in the sample refer to the same risk free rate (Euro area) and the duration bucket are small enough to prevent big differences between D_i and $\overline{D} = 1/J \sum_J D_j$.

- Convexity: $\frac{1}{2}(C_i\sigma_i^2 - \frac{1}{J}\sum_{j=1}^{J} C_j\sigma_j^2)$ there are no missing information by setting to 0 the difference between two small quantities.

- Relative Asset Swap Change: $D_i\Delta_{ASW_i} - \frac{1}{J}\sum_{j=1}^{J} D_j\Delta_{ASW_j}$ when we consider small duration buckets, D_i and D_j don't amplify differences coming from different change in ASW. if the two were identical, we would end up without loss of information with $e\Delta_{ASW} = \Delta_{ASW} - \overline{\Delta_{ASW}}$. Again, it is possible to apply this adjustment given by the small bucket section we study.

It's then possible to rewrite the excess return Equation (3.11) as

$$er_{B_i,K} \approx e_{y,i,t} - D_i * e\Delta_{ASW_i} + o() \tag{3.12}$$

This means that is it possible to express the bond excess return with respect to its duration bucket benchmark by subtracting from the excess yield to maturity the idiosyncratic change in ASW, $e\Delta_{ASW}$ multiplied by the duration.
The former element, excess yield to maturity $e_{y,i,t}$, is straightforward to understand: a bond with a larger carry compared to the benchmark, has an advantage when riskiness doesn't vary.
The latter element is trickier: a positive $e\Delta_{ASW,i}$ implies that change in bond i ASW, $\Delta_{ASW,i}$, is higher then the average change in ASW of bonds in the same duration bucket, $\overline{\Delta_{ASW}}$:

$$e\Delta_{ASW,i} > 0 \qquad \rightarrow \qquad \Delta_{ASW,i} > \Delta_{ASW,b}$$

the bond is becoming relative riskier compared to its benchmark. Given the minus sign before the duration component in Equation (3.12), a positive $e\Delta_{ASW}$ affects negatively the bond excess return. A positive $e\Delta_{ASW,i}$ could result also in a situation in which $\Delta_{ASW,i} < 0$ and this means that the bond is reducing its risk profile less then the average bond. The opposite case arises when $e\Delta_{ASW,i} < 0$, that implies that the bond's risk profile diminishes more compared to the benchmark. From an information perspective, if $er_{B,i,K,t}$ is the random variable of interest, defined on the classical probability triple (Ω, F, P), the first part, $ey_{i,t}dt$ is defined by the filtration at time t, and hence known.

The definition of $e\Delta_{ASW}$ moreover is consistent with the definition of market invariant proposed by Meucci (2005) for the fixed income market. A sufficient theoretical condition is that $e\Delta_{ASW}$ can be defined as a function of total return, that we showed in Equation (3.2). Regarding the necessary condition that requires the risk driver's scatter-plot against its lagged time series to resemble a circular cloud, we show in section (3.5.1) its practical relevance. In section 3.5.1 we test whether Equation(3.12) is correctly specified to derive excess returns.

3.3 Forecasting Procedure

We now turn the attention to handling the projection of the risk driver, to correctly specify bond excess total return. The forecasting procedure we employ stems from the result in the previous section and rather than a time series analysis, this exercise is configured in a panel cross-sectional style. The purpose of the forecasting procedure is to derive a link between bond and equity-specific characteristics available at the time of the choice with future excess change in ASW. The problem can be express as follow

$$\begin{aligned} e\Delta_{ASW_i} &= E[e\Delta_{ASW_i}] + \epsilon \\ E[e\Delta_{ASW_i}] &= f(X_{i,t-1}) \end{aligned} \tag{3.13}$$

where $X_{i,t-1}$ represents the set of features employed.

The corporate bond market is characterized by a complex non-linear structure(Hong et al. (2012)) and, to maximize the explainability, we aim to consider for $f(X_i)$ an algorithm that can extract the non-linearity, subspace characteristics, and the peculiarity among the different sub-samples available. For such a purpose, among statistical learning techniques, Random Forest is the one that fits this problem. We start by analyzing random forest[5] that are extension of Regression Trees. A Regression Tree represents a simple regression technique that splits the parameter space in M sub-spaces $R_1, R_2, ...R_M$ and associates each region a response to match the dependent variable. The sub-regions are chosen by minimizing a certain loss function. The mapping function of a single regression tree, T, can be expressed as

$$T(x; \Theta_t) = \sum_{m=1}^{M} c_m I(x \in R_m | \Theta_t) \tag{3.14}$$

where Θ_t represents the set of hyper-parameters characterizing the tree's properties, namely the sequence of splitting variables, cut-points at each node, and the number of nodes. For each sub-region R_m defined by the algorithm, c_m is computed by taking the mean of all samples in that sub-region. Each sub-space is computed sequentially choosing a splitting variable, j, and the splitting point by minimizing a loss function or an impurity coefficient, such as the Gini Index. The resulting performance is highly affected by the sequential choice of j, given that it is not computationally feasible to search at each step the best split over all the variables.
The main drawback of single trees is that they are extremely noisy and there are two natural ensemble method expansions to address it, namely bagging and boosting. The former steam from bootstrap aggregation, a technique that is capable of reducing variance and is suited for low-bias high-variance frameworks, exactly as trees, and its upgrade proposed by Breiman (2001) that derives the response by averaging out B de-correlated trees. The resulting predictor, following Friedman (2017) notation,

[5]see eg. Brieman et al. (1984) Breiman (2001), Liaw et al. (2002)

has the following form

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T(x;\Theta_b) \tag{3.15}$$

where $T(x;\Theta_b)$ represents the single regression tree. The de-correlation is achieved by choosing for each tree in the forest a sub-set of variables to analyze and by changing the order of the splitting sequence. The final step is to optimize hyper-parameters, Θ_t, by minimizing the Root Mean Square Error (RMSE) function.

The following step is to set the features to employ in the random forest algorithm. Various characteristics have been found to directly impact corporate bond returns. Regarding issuer's characteristics such as credit riskiness (Greenwood and Hanson (2013)) or exposure to factor models (Israel et al. (2017)) are a first driver. Moreover, frictions in the financial market prevent the news to spread quickly and it is possible to observe an asynchronous relationship between stock and bond returns (Cao et al. (2017), Tsai (2014)) as well as in the volatility (Campbell and Taksler (2003)). Liquidity is also regarded as primary factor (Lin et al. (2011)) considering that it is embedded in the credit spread component (Dick-Nielsen et al. (2012)). Given that each attribute analyzed in the literature contributes to the final forecast, we aim to employ both stock information at the issuer level and bond features. We pull together data from the issuer and past data for a bond, and for such a reason, each bond could face three different scenarios:

- No bond Characteristics data: at time t, a bond issued at time $t - h_1$ won't have characteristic computed from time $t - h_2$, with $h_2 > h_1$. Note that, $\forall h_2, \quad \exists h_1 \; s.t. \; h_1 < h_2$ that implies the natural presence of a fraction of bond without historical data. The larger h_2, (the more backward information we retrieve), the greater this percentage.

- No data from Issuer. If an issuer is not publicly listed, there is no measure related to the equity side.

- No measure for bond or issuer. This happens for bonds recently issued from an issuer not listed.

The usual approach is to pool together the data and then either one tries to retrieve missing data employing available ones (Gleason and Staelin (1975), Dixon (1979)) or dismiss the observation with missing values Rubin (1976). In this case, both approaches would present a severe drawback. Filling missing values with available ones would end up averaging out the missing value to the existing one, and would be a mere guess. Deleting all missing values would reduce the size of the bond we would be able to analyze, thus diminishing the possibility to find the bond that outperforms the benchmark. The solution we configure is in the spirit of aggregating multiple forecasters based on different sets of predictors, an extension of the procedure proposed by Rapach et al. (2010) or Lin et al. (2018). In this setting, we build two forecasters:

$$f_x : e\Delta_{ASW,i,t} = rg(X_i, t, 1)$$

$$f_y : e\Delta_{ASW,i,t} = rg(Y_i, t, 1)$$

where the difference between the two forecasters is the set of predictors employed. In particular, we establish one model with bond characteristics forecaster, f_b and the other with equity characteristics measures, f_e. The quest is then to aggregate the two models for the items that present at least one forecast available. Regarding the last case, among the several approaches faced, we select three main possibilities:

- Mean combination forecast (MC): the outcome is the simple average of the two predictors. This is usually a naive benchmark

$$E\Delta_{ASW,i,t} = \frac{f_e + f_b}{2} \tag{3.16}$$

- Bayesian style harmonization (Bates and Granger (1969)): we weight the forecast of each regressor by its confidence, namely the standard deviation of the

forecaster that is easily available from the regressors employed.

$$E\Delta_{ASW,i,t} = \frac{f_e/\sigma_b + f_b/\sigma_e}{(\sigma_e + \sigma_b)} \tag{3.17}$$

- Second optimization framework: we post process f_e and f_b according to macroeconomics variable to gain improvement on the model.

Finally, we take into account the situations in which just one measure is available, either because the bond has been recently issued or the issuer is not publicly listed. We address those cases by employing the available indicator, with the possible extension to consider a penalization term steaming from the only forecaster.

3.4 Data

3.4.1 Financial Data

We arrange the data employed into three main categories, namely bond, equity, and macroeconomic. We obtain from ICE the monthly components of BoFA EG00 index starting in 2010 and retrieve daily data for the bonds. Index components are updated at the beginning of each month. We compute monthly total return for each bond with the standard definition:

$$tr = \frac{P_t + Acc\_Int_t + C_{t-1}^t}{P_{t-1} + Acc\_Int_{t-1}} \tag{3.18}$$

where P is the clean price, Acc_Int is the accrued interest and C is the stream of payments received by bondholders between $t-1$ and t. We consider $t-1$ as the last day of the previous month if the instrument was already in the index, or the first day of the current month if the bond was just included. In such a way, we don't lose information discarding one trading day in each month and the accrued interest chunk, nor do we discard bonds that just entered the index. We then accurately reproduce index performances on a semester basis, that is the ultimate outlook of our research. The database is composed of bonds that have the average rating with

an Investment grade level (BBB- or higher) and we select only bonds that are senior[6]. The resulting database is composed of 5675 unique emissions over 10 years of history (20 semesters) and the total observation amount at 39660.

Table 3.1 shows bonds return distribution characteristics over time. It is possible to appreciate the stylized facts of bond returns, namely the positive skewness associated with positive period and increased dispersion. Observing the distribution of total returns in negative periods (eg: June 2015) it's possible to note that even choosing the bonds in the 90^{th} percentile gives as result a mere 1.5% excess performance compared to the median bond. In a positive period, eg. Jun 19, the bonds in the same position returns a 5% excess return. An effective strategy is then to minimize losses in the negative period to maximize gains in positive ones.

Each semester, we split the bonds universe into 7 duration buckets, according to the bond duration at the beginning of the semester. Table 3.2 shows the duration bucket characteristics as well as the mean performance of each bucket in the various semesters. There are essentially 3 structures in the returns distribution. The first one is pictured by long-duration bonds as best performer, the second in which mid-duration buckets overperform, and the third in which the best performers are at short duration horizon. These differences are linked to the movements of the curve and the amplification mechanism from duration.

From the EG00 database, we extract quantitative and qualitative bond characteristics implemented as features. We shortly describe some of the main data we employ and their transformation procedure.

From price elaboration, we extract past returns over different periods as well as risky characteristics, namely standard deviation and downside deviation,

We employ also duration characteristics that indicate to the algorithm the location of the bond in the curve, credit risk component of bond such as ASW, that is the driver

[6]And thus don't display equity style characteristics.

Table 3.1. Universe Statistics and summary

	Index	Return Stats				Return Quantiles						
	count	mean	std	min	max	5%	10%	25%	50%	75%	90%	95%
Jun-10	1425	4.3%	3.2%	-10.0%	23.7%	-0.2%	1.3%	2.7%	4.3%	5.9%	7.6%	9.1%
Dec-10	1591	1.4%	3.7%	-30.5%	29.4%	-2.3%	-1.0%	0.2%	1.0%	1.9%	3.8%	6.9%
Jun-11	1607	1.3%	2.4%	-16.6%	16.4%	-0.8%	-0.2%	0.2%	0.8%	1.9%	3.8%	5.2%
Dec-11	1655	0.4%	7.3%	-38.4%	17.5%	-15.2%	-8.6%	-0.8%	2.3%	4.4%	6.4%	7.6%
Jun-12	1588	5.5%	4.7%	-11.7%	35.3%	0.4%	1.5%	2.7%	4.8%	7.6%	11.0%	13.6%
Dec-12	1571	6.9%	6.0%	0.6%	50.1%	1.3%	1.6%	2.9%	5.1%	8.5%	14.6%	19.8%
Jun-13	1552	0.6%	2.0%	-6.0%	31.7%	-1.5%	-0.9%	-0.2%	0.4%	1.0%	2.1%	3.2%
Dec-13	1593	2.2%	2.2%	-5.3%	17.9%	0.3%	0.5%	0.9%	1.5%	2.8%	5.0%	6.8%
Jun-14	1635	5.0%	3.2%	-1.0%	22.5%	0.9%	1.1%	2.2%	4.7%	7.2%	9.2%	10.6%
Dec-14	1750	3.2%	3.5%	-24.2%	17.1%	0.5%	0.7%	1.2%	2.5%	4.8%	7.4%	8.9%
Jun-15	1761	-1.0%	2.0%	-14.4%	5.8%	-4.5%	-3.5%	-2.0%	-0.4%	0.2%	0.6%	1.0%
Dec-15	1981	1.0%	3.3%	-33.8%	8.8%	-1.8%	0.0%	0.5%	1.2%	2.2%	3.4%	4.1%
Jun-16	1989	4.1%	4.0%	-4.6%	34.1%	0.3%	0.6%	1.3%	3.1%	5.8%	8.5%	11.0%
Dec-16	2172	0.6%	2.1%	-8.8%	17.0%	-2.1%	-1.2%	-0.2%	0.2%	1.1%	2.7%	4.4%
Jun-17	2265	0.7%	1.9%	-5.2%	13.6%	-0.8%	-0.5%	-0.2%	0.2%	0.9%	2.7%	4.9%
Dec-17	2443	1.8%	1.6%	-3.6%	14.5%	0.2%	0.3%	0.7%	1.4%	2.3%	3.7%	5.1%
Jun-18	2538	-0.4%	1.6%	-10.4%	9.7%	-3.7%	-2.1%	-0.6%	0.0%	0.3%	0.7%	1.0%
Dec-18	2668	-0.6%	1.6%	-17.1%	3.9%	-3.0%	-2.0%	-0.9%	-0.2%	0.1%	0.5%	0.9%
Jun-19	2777	5.1%	3.9%	-1.0%	22.8%	0.4%	0.7%	1.9%	4.3%	7.5%	10.5%	12.6%
Dec-19	2935	0.5%	1.4%	-5.0%	11.0%	-0.7%	-0.5%	-0.2%	0.1%	0.9%	2.0%	3.0%
Jun-20	3098	-1.5%	2.2%	-23.4%	10.1%	-5.5%	-4.0%	-2.2%	-1.0%	-0.3%	0.2%	0.7%
Dec-20	2858	3.8%	3.1%	-5.4%	21.2%	0.6%	0.8%	1.5%	2.9%	5.2%	7.8%	10.1%

Table (3.1) shows the main properties for the Index considered, BofA ML Euro Corporate IG, ER00. Considered bond are those with seniority higher then "Subordinated". Each rows refers to the semester ending in the last day of the month indicated, hence Jun 2013 is the period from 1 Jan 2013 until 30 Jun 2013. The mean and std(Standard Deviation) value are computed equally weighting the issues in the portfolio. The quintiles columns show the distribution for total returns each semester.

Table 3.2. Universe performance for duration buckets

	Duration Bucket (Y)						
	0-1	1-2	2-3	3-5	5-7	7-10	10+
Jun-12	3.40%	3.76%	5.18%	6.46%	7.19%	7.15%	9.46%
Dec-12	2.84%	3.86%	6.96%	8.10%	8.15%	7.05%	14.25%
Jun-13	1.00%	0.85%	0.56%	-0.10%	-0.26%	-0.05%	3.95%
Dec-13	1.10%	1.78%	2.33%	2.14%	2.08%	1.22%	0.88%
Jun-14	1.13%	1.70%	3.39%	5.77%	7.37%	6.09%	11.69%
Dec-14	0.70%	1.01%	1.64%	3.38%	5.10%	6.07%	10.48%
Jun-15	0.33%	0.30%	-0.10%	-1.16%	-2.27%	-2.44%	-3.78%
Dec-15	0.38%	0.39%	0.84%	0.77%	1.61%	1.79%	2.44%
Jun-16	0.51%	1.18%	2.15%	3.77%	5.05%	6.01%	12.09%
Dec-16	0.30%	0.61%	1.02%	0.72%	0.78%	-0.62%	-3.03%
Jun-17	0.21%	0.21%	0.51%	0.82%	1.56%	0.64%	-0.12%
Dec-17	0.24%	0.47%	1.15%	1.98%	2.87%	2.51%	2.97%
Jun-18	-0.08%	0.00%	-0.20%	-0.70%	-1.14%	-0.57%	-0.18%
Dec-18	-0.06%	-0.24%	-0.43%	-0.67%	-0.76%	-0.83%	-1.22%
Jun-19	0.76%	1.40%	3.16%	5.59%	7.84%	7.77%	13.13%
Dec-19	0.07%	0.14%	0.29%	0.60%	0.89%	0.76%	1.45%
Jun-20	-0.60%	-0.96%	-1.50%	-1.79%	-1.75%	-1.49%	-1.55%
Dec-20	1.12%	1.78%	2.78%	4.24%	5.25%	7.17%	11.06%

Table (3.2) shows the performance of BofA ML Euro corporate bond index, EG00, splitted for duration bucket. At the beginnin of eacn each semester, bond are clustered in duration buckets according to the duration they exhibit. The return is computed by equally weighting the issues in the portfolio. For each duration bucket K, the columns header present $[min_k\text{-}max_k)$.

of interest, and pure interest measured as YTM. We utilize also Amount Outstanding as a proxy for liquidity. Other factors are computed by combining this measure, such as yield to maturity(YTM) over Price that indicates the put to the par component of each bond relative to the emission level and Sharpe ratio. All the quantitative measures are employed as stock and flow at different time horizons back, to derive the movement. In addition, we translate qualitative measures such as Rating into features that can be exploited from the algorithm.
Statistical Learning algorithms improve forecasts especially when standardized values are passed through the algorithm as features. In this context, we standardize in a cross-sectional way rather than a time-series, given the model characteristics, and within each duration bucket. The feature that the algorithm receives to forecast Idiosyncratic change in ASW is then the relative position within that duration bucket, Table shows the main statistics of the principal measures employed.

The following part of the analysis deal with firm specification, to extract issuer characteristic. We link each bond to its issuer following the unique issuer code identifier retrieved by Reuters and the issuer information. For equity characteristics, we utilize month-on-month measures that allow us to disentangle firm-specific characteristics. Then we employ issuer data to forecast bond performances by considering equity measures retrieved from price information such as percentage change, volatility as well as technical indicators such as the Moving Average Convergence Divergence (MACD) and the Relative Strength Index (RSI) and volume information. Also from issuer data, we retain qualitative information, namely sector and state, that transform into numbers to correctly segment the space.

Table 3.4 shows the main statistics of the principal measures employed. We standardize just in the mean also the measures here employed. When we were to take a neural network algorithm, a complete standardization with the variance would be needed.

Table 3.3. Bond features statistics and summary

	Panel (A) Mean											panel (B) Standard Deviation										
	'10	**'11**	**'12**	**'13**	**'14**	**'15**	**'16**	**'17**	**'18**	**'19**	**'20**	**'10**	**'11**	**'12**	**'13**	**'14**	**'15**	**'16**	**'17**	**'18**	**'19**	**'20**
ret_f	4.3%	0.8%	3.4%	2.9%	3.1%	1.1%	2.6%	0.2%	0.7%	2.0%	-0.5%	2.75%	1.86%	3.94%	2.93%	2.50%	2.93%	3.77%	1.33%	1.17%	2.75%	1.85%
ret_fe	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.30%	1.57%	3.80%	2.72%	1.55%	1.14%	2.66%	1.29%	0.94%	1.80%	1.56%
oas_f	152.7	156.6	253.2	133.3	97.1	96.4	119.5	99.7	93.6	121.8	115.9	83.2	97.4	150.6	74.9	49.6	63.8	78.6	36.5	36.5	58.8	57.4
oas_fe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	76.71	69.74	141.43	71.58	41.99	39.18	61.42	35.43	33.86	55.70	51.10
ytm_f	3.36	3.49	3.18	1.89	1.66	1.13	0.98	0.75	0.76	0.78	0.61	1.07	1.19	1.61	0.97	0.81	0.81	0.90	0.62	0.63	0.71	0.62
ytm_fe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.74	1.43	0.75	0.45	0.41	0.62	0.38	0.36	0.57	0.51
gain_f	3.2%	3.4%	3.2%	1.8%	1.6%	1.1%	1.0%	0.7%	0.7%	0.8%	0.6%	1.2%	1.4%	1.9%	1.0%	0.8%	0.9%	1.0%	0.6%	0.6%	0.7%	0.6%
gain_fe	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	0.9%	1.8%	0.8%	0.5%	0.4%	0.7%	0.4%	0.4%	0.6%	0.5%
dur_f	4.2	4.1	3.9	4.3	4.6	5.0	5.1	5.3	5.2	5.1	5.1	2.25	2.35	2.38	2.59	2.65	2.94	2.95	2.97	2.97	2.94	3.25
std_f	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.04	0.03	0.04	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.07
std_fe	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.03	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.02	0.04
p_f	105.0	103.1	104.2	107.9	107.6	108.6	107.3	106.2	104.8	104.1	104.5	6.14	5.99	7.48	7.00	6.85	8.58	8.94	7.91	7.08	6.62	6.90
p_e	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.07	5.77	7.62	7.02	6.88	8.23	8.75	8.01	7.14	6.42	6.83
sharpe_r	1.7	0.5	1.4	1.6	1.8	1.0	1.2	0.2	0.4	0.2	0.0	1.35	0.64	1.08	0.97	0.81	0.67	0.62	0.57	0.39	0.32	0.12
ret_f_2m	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
fwd	2.1%	2.6%	1.43%	0.99%	1.01%	0.51%	0.24%	0.32%	0.41%	0.08%	-0.2%	0.4%	0.45%	0.36%	0.42%	0.46%	0.29%	0.31%	0.33%	0.36%	0.28%	0.1%
rf	1.9%	2.4%	1.4%	0.8%	0.9%	0.4%	0.2%	0.2%	0.3%	0.0%	-0.2%	0.5%	0.5%	0.3%	0.4%	0.4%	0.3%	0.3%	0.3%	0.4%	0.3%	0.1%
diff	0.2%	0.2%	0.1%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Table (3.3) shows the mean and the standard deviation for the entire period (2010-2020) of the dataset employed as features for the random forest regression algorithm. variable that display 0 mean are standardized.

Table 3.4. Summary Statistics equity

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
ret	0.221	-0.097	0.216	0.051	0.068	-0.005	0.156	0.087	-0.109	0.094	0.169
dvd	3.14	3.74	3.83	3.51	3.04	3.13	3.32	3.13	3.28	3.56	3.02
mkt_cap	9.10	9.17	9.18	9.28	9.34	9.34	9.32	9.40	9.43	9.42	9.40
book_v	13.99	14.55	14.45	14.99	15.69	16.11	15.69	15.91	16.75	17.14	17.82
roe	7.17	7.79	7.77	8.22	6.87	3.48	3.60	5.14	3.96	0.92	-4.59
longdebt_a	22.07	21.66	22.62	23.63	23.12	24.31	25.48	25.60	25.32	25.51	24.85
totdet_a	25.76	25.06	26.18	27.04	26.17	27.22	28.34	28.80	28.31	28.37	27.72
capex	4.13	4.99	4.92	4.77	4.74	4.42	4.08	4.10	4.24	4.02	3.02
ppe	8.25	8.26	8.26	8.26	8.24	8.24	8.25	8.26	8.27	8.28	8.39
mom	2.09	0.58	1.48	3.03	0.89	0.45	1.16	1.68	0.27	1.03	1.94

Table (3.4) shows the main statistics for the equity features employed in the statistical learning algorithm. From bond ISIN it is possible to retrieve Issuer ID and its unique identifier. We collect fundamental data and the average is done considering the number of firm.

3.4.2 Climate Risk Data

The data we employ to measure climate risk is related to the transition risk component. From Datastream, we obtain Scope 1 and Scope 2 Emission at issuer levels from 2009 until the end of the sample. We consider the log of this value. Employing data obtained for stock fundamentals and financial data, we derive standardized emission by sales. In addition, we compute the percentage change of carbon emission. The frequency of the emission data is yearly, which creates a mismatch with the frequency of the bond dataframe. We however consider the same measure for the two-semester periods inside each year. There are two years, 2019 and 2020 in which the amount of Carbon emission diminishes but the percentage change doesn't reflect this variation. This is due to some firms that left the database in that year, so the value of carbon emission for that year is not available.

3.5 Results

3.5.1 Excess return decomposition

In section 3.5.1 we derive $e\Delta_{ASW}$, the market invariant for corporate bond excess return. In this section we show that the specification of Equation (3.12) is correctly defined and able to capture the inherent driver for excess return, employing the dataset from 2010 until 2020 with semester frequency.

We firstly assess the correctness of interpreting $e\Delta_{ASW,t}$ as the risk driver, following

Table 3.5. Transition Risk summary

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
s1	7.2E+06	7.1E+06	6.9E+06	6.6E+06	6.6E+06	6.3E+06	6.1E+06	5.9E+06	6.4E+06	5.0E+06	4.4E+06
s2	1.1E+06	1.1E+06	1.1E+06	9.9E+05	1.0E+06	9.2E+05	9.0E+05	8.3E+05	8.9E+05	7.5E+05	6.6E+05
ds1	-1.6%	2.8%	0.2%	0.9%	3.0%	-1.8%	1.0%	-0.4%	3.1%	2.7%	-0.1%
ds2	2.3%	4.0%	0.5%	-2.6%	-2.6%	-2.1%	1.5%	-1.9%	4.3%	8.3%	-0.5%

Table (3.5) shows the summary statistics for the emission data at firm level. s1 is Scope1, s2 represent Scope2, ds1 is the yearly percentage change in Scope1 emission and ds2 is the yearly percentage change in Scope2 emissions. Issuer are selected starting from the BoFA ML Euro corporate IG. The table shows the yearly average equi-weighting the firms

the definition of market invariant proposed by Meucci (2005). In Figure (3.1) we plot the excess change in Asset Swap Spread against its lagged value. It 's possible to note that the shape of the cloud is circular and centered in 0, so we can safely conclude that the characterization of $e\Delta_{ASW}$ is a market invariant.

Figure 3.1. Scatter plot with lags for $e\Delta_{ASW}$

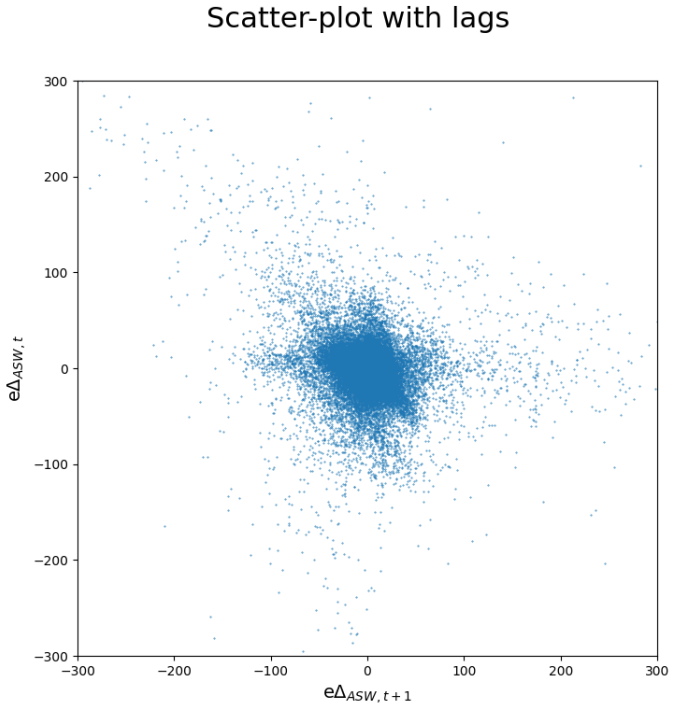


Figure (3.1). Scatter plot that shows the relationship between lags of excess change in Asset Swap Spread, $e\Delta_{ASW}$ from time t, placed in the x axis, to time $t+1$, placed in the y-axis, where the time-frame is a semester. To correctly define a market invariant, the shape of the scatter-plot must resemble a circular cloud. Sample period is 2010-2020 for BofA ML EG00.

The second valuation regards the precision of Equation (3.12). We handle several approximations along the derivation, such as the drop of risk free rate component, the average of change in Asset Swap Spread over different point of the yield curve, and the discard of higher derivative moments such as convexity. We demonstrate that those approximations don't undermine the exploitative power of $e\Delta_{ASW}$ over the excess return. For each semester, we recover the bond's effective excess return er_t and compare to the excess return computed applying Equation(3.12), employing

the excess yield to maturity, ey_{t-1} and the measured idiosyncratic change in Asset Swap, $e\Delta_{ASW}$.

With regard to the length of the duration bucket, the larger the bucket, the higher the possibility of a miscalculation in the terms that encompass the duration, due to its amplification properties. The shorter the bucket length, the less bond in each bucket to analyze and the more imprecise computation of the average change in the ASW, $\overline{\Delta}_{ASW}$. We set the duration bucket length equal to one year [7]. We test in particular:

$$er_t = \alpha + \beta * e\tilde{r}_t + \epsilon \tag{3.19}$$

where $e\tilde{r}_{i,t} = ey_{i,t-1} - D_i * e\Delta_{ASW,i}$ refers to Equation 3.12.

Figure 3.2. Estimation of excess return through idiosyncratic asset swap change

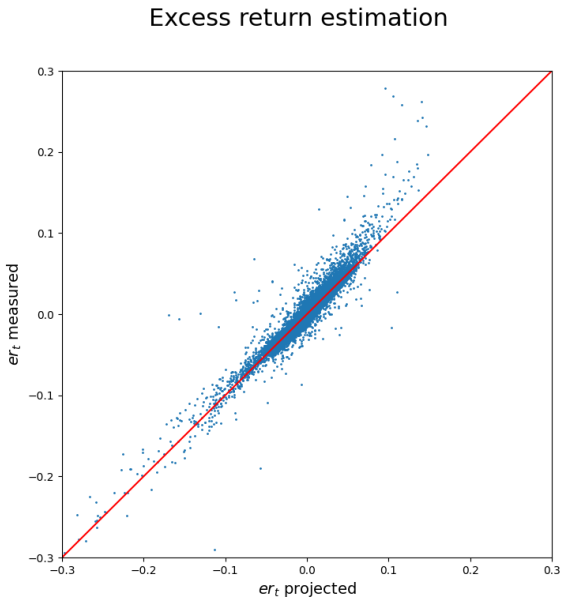


Figure (3.2). Scatter plot that shows the relationship between estimated excess returns, $\tilde{er}$ through the decomposition of idiosyncratic change in ASW and realized excess return, er. On x axis, er, the realized excess return. On Y axis, $\tilde{er}er$, the forecasted excess return. The red line shows the perfect relationship, with the equation $y = x$. Sample period is 2010-2020 for BofA ML EG00.

[7] We explore duration buckets length from 5 years to 0.3 years, result presented in Table 3.6. As expected, from 5 to 2 years the duration multiplication error is high. From 1 year to 0.5 years duration length, the results still get better but not in a significant way and there's no gain in shrinking the bucket

Table (3.7) shows the estimation result for Equation (3.19) through Ordinary Least Square (OLS). The first evidence is that α is statistically not different from 0, that allows to assess that there are no constant effect missing from the formulation. In addition, β parameter is 1.01[8], that implies that Equation (3.12) exactly characterizes the excess return dynamics. Considering the R^2 coefficient, it is above 90% in the total sample and the only duration bucket in which is lower is the 10+ bucket. This approach is thus able to disentangle the great majority of the variance in the sample.

In Figure 3.2 we plot the actual and estimated excess return, on the y and x axis respectively, and the straight line $y = x$ where there is perfect match between the two. It is possible to note that all the points are close to the linear relationship, with just some exception. In particular, those exception are bonds with high duration that are the ones difficult to characterize.
This result confirms what moves the analysis: by extracting excess return, in addition to forget all higher moments, one does not have to reprice all the bond risky characteristics, it just have to forecast the relative movement with respect to the index. From a slightly different prospective, among a certain amount of firm with different characteristics, one does need to highlight who will become more or less risky compared to the others.

We investigate whether there are differences accounting for duration bucket. In this way we can test whether the algorithm produces bias at certain duration buckets. Table (3.7) report the result of the estimation of Equation 3.19 in 4 different duration buckets, namely smaller then 3 years, between 3 and 5 years, between 5 and 7 years and greater then 7 years.

It's possible to see that the intercept coefficient is null in all the buckets, Equation

[8]We run a F-test on the β coefficient to test $H_0 : \beta = 1$ and we do not reject the null hypothesis.

Table 3.6. Duration lenght evidence for modelling excess return

	5Y	4Y	3Y	2Y	1Y	0.5 Y
$\hat{er}$	1.007	1.006	1.011	1.013	1.021	1.013
	***	***	***	***	***	***
α	-1e-05	4e-05	7e-07	-3e-06	-1e-05	-2e-05
	(-0.2390)	(0.9538)	(0.0180)	(-0.1094)	(-0.3346)	(-0.8621)
No. Obs.	39660	39622	39686	39622	33355	36809
Cov. Est.	Clust	Clust.	Clust.	Clust.	Clust.	Clust.
R-squared	0.7960	0.8287	0.8586	0.8875	0.9032	0.9091

T-stats reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (3.6) shows the same OLS estimation for $E_{tr} = \alpha + \beta \hat{E_{tr}} + \epsilon$ where$\hat{e_{tr}} = ey - D * E\Delta_{ASW}$ varying the length of the duration bucket employed to compute $E_{\Delta ASW}$, that are indicated in the columns. 5Y stands for the duration buckets: [0, 5); [5, 10); [10+. 4Y stands for[0, 4); [4, 8); [8, 12); [12+. 3 is [0, 3); [3, 6); [6, 9); [9, 12) [12+). 2, 1, and 0.5 year the same up to 12 year and the last bucket is [12 +. Sample Period is 2010-2020.

Table 3.7. Excess Return and model specification - estimation

	<3Y	3-5Y	5-7Y	7-10Y	10Y+	Total
$\hat{er}$	0.89	0.99	1.04	1.09	1.01	1.01
	***	***	***	***	***	***
α	-7e-05*	-2e-05	2e-06	-9e-05	-2e-03	-2e-05
	(-1.72)	(-0.52)	(0.03)	(-1.01)	(0.83)	(-0.86)
Cov. Est.	Clust.	Clust.	Clust.	Clust.	Clust.	Clust.
R-squared	0.9007	0.9367	0.9371	0.9318	0.7726	0.9091

T-stats reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (3.7) shows the OLS estimation for $E_{tr} = \alpha + \beta \hat{e_{tr}} + \epsilon$ where $\hat{e_{tr}} = ey - D * E\Delta_{ASW}$ with duration bucket of 1 year estimated in the different part of the duration region. Bond are assigned in each region according to the value of the duration at the beginning of the period. Sample Period is 2010-2020 for Euro IG ER00.

Table 3.8. Walk Forward - ASW forecast

Panel (A): Percentage of analysis

	M1	M2	Pooled MC	Pooled BC
Avg Analyzed	65.55%	79.53%	86.25%	86.25%
Min	60.13%	68.12%	79.53%	79.53%
Max	71.17%	85.28%	91.25%	91.25%

Panel (B): Summary Statistics forecasting excercise

	M1	M2	Pooled MC	Pooled BC
β	0.78	0.82	0.83	0.84
R^2	13.38%	16.82%	24.51%	25.78%
Accuracy	55.6%	53.49%	57.42%	58.12%
RMSE	0.92	0.89	0.85	0.84

Table(3.8) shows the results from the walk-forward exercise from 2016 until 2019 with a rolling training period of 12 semesters. M1 stands for the model with bond characteristics as features, M2 for the Equity characteristic as features. $Pooled_MC$ is the MC pooling algorithm of M1 and M2 and $Pooled_BC$ is the BC pooling algorithm of M1 and M2. The average analyzed is the simple average of the 9 semester in out-of-sample exercise. β is the estimated coefficient of $E\Delta_{ASW} = \beta\overline{E\Delta}_{ASW} + \epsilon$. Accuracy is the percentage of time $E\Delta_{ASW}$ and $\overline{E\Delta}_{ASW}$ display the same sign. RMSE is computed as fraction of the RMSE of a random classifier that sets all $\overline{E\Delta}_{ASW}$ equal to 0

(3.19) is not posing a bias in any duration bucket. Regarding the β coefficient, it's possible to note that in duration bucket 3-5 and 5-7 and 10+ it's very close to 1, the target value. In the shortest duration bucket the coefficient is approximately 0.9, implying that the effective excess return is sightly smaller then the one we are able to forecast. It the duration between 7 and 10 however it happens the opposite, with a coefficient of 1.1, that means a forecast that reduces the differences of the excess return. When we look at R^2, it's above 90% for all the buckets with except for the longer duration where is falls at 77%. This is due to the small number of bonds in each duration bucket, especially for long duration for which it is not feasible to build 1 year duration bucket.

3.5.2 ASW change forecast

We establish a correct way to disentangle excess return through the asset swap spread. This element is the crucial to be forecasted to obtain a correct specification for the excess return evaluation. We implement the forecasting procedure described in Section 3.3) and test its performance in projecting at the horizon the risk driver. The aim is to assess whether the random forest algorithm is able to forecast excess change in asset swap and the enhancement obtainable through the pooling mechanism. Considering the dataset available and the models described, we test the following specification.

- M1, that indicate the exercise in which the forecasted value is $E\Delta_{ASW_i}$ through the bond characteristic features
- M2: that indicates the exercise in which the forecasted value is $E\Delta_{ASW_i}$ through the equity characteristic features
- Pooled MC: that indicates the exercise in which we pool the results of M1 and M2 with MC algorithm
- Pooled BC: that indicates the exercise in which we pool the results of M1 and M2 with Bayesian weighting algorithm

We run a walk forward exercise, with a rolling estimation window of 12 semesters, starting in December 2016 until December 2020, and obtain 9 out-of-sample period. Table 3.8 shows the results by averaging the out-of-sample periods. The first thing to notice is the gain in analysis extent that the pooling mechanism allows to make. Bond and Equity features alone are able to predict between 65% and 85% of the sample, given missing values[9]. By pooling the two results, we obtain a coverage that is in average at 86% with a maximum value of 91%. [10] We compare the model output against the realized value. The models on average manage to find the exact direction

[9] given that the algorithm handles paste returns, it is not possible to analyze bond issued in the previous months.

[10] The percentages are equal between the two pooling methods because the percentage is not affected by the method itself, given that they don't present any stricter requirement with the respect to the original models to analyze each bond.

55% of the times, with a partial increase in the pooled mechanism to 58%. Moreover, with respect to a random forecaster that assigns to each bond no change in ASW, the average RMSE diminishes by roughly 10%.

Finally, we analyze the variables the random forest algorithm regards as the most important to explain changes in ASW in the two different subsets. Table (3.9) shows the top three variables the random forest employs to address idiosyncratic change in asset swap spread. We point out that we are not assessing whether the variable has a positive or negative impact, but rather that its values are employed by the random forest regressor as primary driver.

Regarding Bond characteristics, we find that the excess Price, ($ExcPrice$) is the first to be chosen, followed by the downside deviation ($Downside$) and Future expected Gain($GainFe$). An important role is caught by the Yield to Maturity and the same Asset Swap elaboration. Regarding the excess price factor, it is similar to the approach developed by Nozawa (2013) that matched corporate to governative bond, here we check against parity. When looking at the downside risk, already documented as a factor in corporate bond total return (Bai et al. (2019); Augustin et al. (2020)), our conjecture regards the fact that higher downside risk is reflected by larger variation in asset swap spread, so the algorithm tends to cluster bond that are likely to move the most and the least.

Regarding Equity characteristics, we find the most important feature to be the issuer stock past total return. This finding is well documented in the literature (Ilmanen (2003); Cappiello et al. (2006)) and in this case indicates that information flow in bond market circulates with a certain delay compared to equity market. When looking at standard deviation, Std, it is either correspondent to the bond downside deviation, and to the effect of idiosyncratic volatility explored by (Campbell and Taksler (2003)). In Campbell and Taksler (2003) work, the connection is established between the idiosyncratic volatility and the cross section of yield change, we furthermore establish the link to the idiosyncratic change in asset swap spread, the specidic driver for corporate yield.

Table 3.9. Variable importance ranking

	M1	M2
1	Exc Price	Ret
2	Downside	Std
3	Gain Fe	Mkt Cap
4	YTM	LongDebt
5	ASW	ROE

Table(3.9) shows the first five elements by ranking importance for the two models estimated. M1 refers to the model built upon Bond characteristics data. M2 stands for the equity characteristic model. Feature importance is derived by considering the Gini coefficient, that ranks features according to the impurity reduction at each node employing the Gini Index(Nembrini et al. (2018))

3.5.3 Portfolio performance

The leading purpose of the analysis is to derive a portfolio that over-performs the related benchmark, in an exercise of bond selection. We analyze the excess return implicitly forecasted by the model through the Equation (3.12)

$$\overline{er}_{i,t} = ey_{i,t-1} - D * e\Delta_{ASW,i,t}. \tag{3.20}$$

where $ey_{i,t-1}$ is the deterministic component.

We emphasize that duration acts as an amplification component and it retrains heteroskedastic characteristics for er. As stated before, the bond universe is split in duration buckets of 1 year, and from the the forecasted ASW, we are left with an overperformance inside a certain bucket.

The main drawback of this derivation is that it is not feasible to compare excess returns of bonds in different duration buckets. In fact, given a bond's a excess return, er_a, associated to a duration bucket d_a, and a bond's b excess return er_b related to different duration bucket d_b, $d_a \neq d_b$, the case $er_a > er_b$ doesn't entail $tr_a > tr_b$. More generally, $er_a \lesseqgtr er_b \quad \not\Rightarrow \quad tr_a \lesseqgtr tr_b$. The an higher er_a refers to a negative benchmark return while er_b applies to a positive benchmark return and thus, $tr_a \leq tr_b$

In our approach this is a minor threat given that in the asset allocation procedure, the

strategic asset allocation and the tactical asset allocation have already contributed to define the share of the portfolio that has to be allocated in the different buckets. In addition, we challenge the position by a general investor without such a report. In this light, the investor could face the situation formerly described, where selecting bond based on $e\Delta_{ASW}$ or er would end up in selecting duration buckets that underperform the benchmark [11] The most simple solution is to invest in each duration bucket a portfolio fraction equal to the share of the bond in the index that are in that duration bucket [12]. Following this procedure, for each duration bucket d_i, the total return of the portfolio, $tr_{p,d_i} > tr_{b,d_i}$ given that the algorithm selects bonds with a positive er in that duration bucket. Then, the portfolio return is the weighted average of the single duration portfolios,

$$tr_P = \sum_{i}^{N_b} w_d * tr_{p,d_i}$$

it's easy to understand that if $\forall i, tr_{p,d_i} > tr_{b,d_i} \Rightarrow tr_P > tr_B$. A more challenging approach, would leave the algorithm to select bonds with an higher expected return, given that this number is not small, we could overcome the previous risk. The risk is that in this way the algorithm choose long duration bonds, easing the way for good excess returns. For the prosecution of the work, we investigate the first choice, having the same percentage of bond in each duration bucket.

We conduct two tests to evaluate the portfolio performances. The former is meant to evaluate the performance of a portfolio of 100 bonds rebalancing the component each semester and the second in order to compare the performance of different portfolios based on the bond percentile.

The first is coherent with an investor logic, that has a lump sum to invest. In this setting, the investor considers the minimum investment threshold of each bond, in

[11] Given as granted the forecasted values, this is the only situation in which this algorithm could produce negative results.

[12] Note that to be as precise as possible, the computation of the share of bond in the duration bucket should follow the same rule employed to compute index total return, most often market weighted

the most case equal to 100.000€, and chooses the number of bond in which to invest according to it's budget. We test the performance of the algorithm by setting a dynamic investment portfolio determined by the first 100 bonds the algorithm rank for excess return. We test the same specifications employed in section(3.5.2), with M1 and M2 respectively the output from the random forest algorithm employed on bond and equity characteristics, Pooled the MC pooling algorithm and Lasso a control algorithm employed as a reference.

Table (3.10) presents the result of a walk-forward exercise starting in June 2016, the first semester ending in December 2016. The nine semesters present an heterogeneous sample for market conditions, encompassing situations of credit risk reduction (2017 and 2019), stressed period (2018), the Covid induced crisis (Jun-2020) and the resilient gain in the second semester(Dec-2020).

Panel A shows the total return in each semester for the benchmark and the different portfolios and we compare the single portfolios against the benchmark. Overall it is possible to note that in positive time, characterized by decrease in credit risk and positive benchmark total return, the algorithm forecasting idiosyncratic change in ASW manage to outperform the benchmark with an the performance higher as more positive is the benchmark return. In down-turning moments, the portfolios generally perform lower then the benchmark, with a discrepancy lower than the one of the positive semesters. In addition, the pandemic crisis offers a good test for the robustness of the different models in very adverse situation. The portfolio percentage loss, compared to benchmark, is between 1.5% and 2%, well below the gains in positive moments that enable an overperformance from 2% to 4%.

We then compare the number of bonds in each semester for each strategy that outperform the benchmark, displayed in Panel D. A higher total return can derive from some bond well outperforming the benchmark and the majority being close or below, or the entire set slightly over-performing the benchmark. We see that, with the exception of the first semester in 2020 characterized by Covid, in all semesters the percentage of bond outperforming the benchmark is well above 50% with just a

Table 3.10. Portfolio statistics - 4 algorithms

Panel A: Total Return

	Nr Bond	Bench.	M1	M2	Pooled	Lasso
Dec - 2016	2033	0.2%	2.4%	2.7%	2.7%	1.4%
Jun - 2017	2157	0.3%	1.7%	1.7%	1.7%	1.1%
Dec - 2017	2293	1.4%	2.6%	1.4%	2.7%	1.9%
Jun - 2018	2412	-0.1%	0.0%	-0.1%	0.0%	-0.1%
Dec - 2018	2503	-0.6%	-1.5%	-0.6%	-1.2%	-1.2%
Jun - 2019	2643	4.5%	7.6%	4.5%	8.7%	6.3%
Dec - 2019	2830	0.3%	2.3%	0.3%	2.3%	1.2%
Jun - 2020	3001	-1.4%	-3.0%	-1.4%	-3.3%	-2.5%
Dec - 2020	2774	3.5%	7.5%	3.5%	8.2%	5.7%

Panel B: Portfolio Performance

	Bench.	M1	M2	Pooled	Lasso
Dec - 2016	100.2	102.4	102.7	102.7	101.4
Jun - 2017	100.5	104.1	104.5	104.5	102.6
Dec - 2017	101.9	106.8	107.2	107.3	104.5
Jun - 2018	101.8	106.8	107.2	107.3	104.4
Dec - 2018	101.2	105.2	105.4	106.1	103.2
Jun - 2019	105.8	113.2	113.7	115.2	109.7
Dec - 2019	106.1	115.8	116.6	117.9	110.9
Jun - 2020	104.6	112.3	113.1	114.0	108.2
Dec - 2020	108.2	120.7	121.9	123.3	114.4

Table(3.10) shows characteristics for the walk forward exercises conducted, from 2016 until 2019 with a rolling training period of 12 semesters. M1 stands for the model with bond characteristics as features, M2 for the Equity characteristic as features and *Pooled* is the BC pooling algorithm of M1 and M2 and LASSO a control model. The portfolios are composed by 100 bonds that represent in each semester the ones each algorithm considers to be the best performer. Panel A shows the Total Return of each portfolio in the semesters while Panel B the evolution of 100 Euros invested in the portfolio.

Table 3.11. Portfolio statistics - 4 algorithms

Panel C: Duration

	Bench.	M1	M2	Pooled	Lasso
Dec - 2016	5.25	5.74	5.75	5.75	4.57
Jun - 2017	5.32	6.24	6.32	6.32	5.09
Dec - 2017	5.40	6.58	6.61	6.31	5.07
Jun - 2018	5.42	6.15	6.19	6.21	5.14
Dec - 2018	5.33	6.16	6.28	6.24	5.02
Jun - 2019	5.29	6.24	6.27	6.35	4.97
Dec - 2019	5.24	6.37	6.49	6.24	4.81
Jun - 2020	5.37	6.26	6.43	6.39	5.03
Dec - 2020	5.21	6.19	6.17	6.13	4.86

Panel D: Percentage beating benchmark

	Bench.	M1	M2	Pooled	Lasso
Dec - 2016	49.1%	87.2%	88.0%	88.0%	83.0%
Jun - 2017	43.6%	83.2%	82.0%	84.0%	76.0%
Dec - 2017	42.0%	70.4%	69.0%	72.0%	53.0%
Jun - 2018	66.5%	61.6%	59.0%	60.0%	61.0%
Dec - 2018	61.3%	57.6%	58.0%	48.0%	59.0%
Jun - 2019	47.3%	66.4%	71.0%	77.0%	60.0%
Dec - 2019	39.5%	84.8%	87.0%	82.0%	66.0%
Jun - 2020	57.9%	33.6%	33.0%	31.0%	40.0%
Dec - 2020	41.5%	82.4%	85.0%	83.0%	68.0%

Table(3.11) shows characteristics for the walk forward exercises conducted, from 2016 until 2019 with a rolling training period of 12 semesters. M1 stands for the model with bond characteristics as features, M2 for the Equity characteristic as features and *Pooled* is the BC pooling algorithm of M1 and M2 and LASSO a control model. The portfolios are composed by 100 bonds that represent in each semester the ones each algorithm considers to be the best performer. Panel C shows the mean duration of each portfolio, computed by equally weighting each bond duration. Panel D present the percentage of bond in each portfolio that outperform the benchmark. The benchmark column is not equal to 50 because the average doesn't match the median and the distribution is skewed.

couple of semester around 60% and some close to 90%.

The following analysis we carry regards the distribution of bond returns in the pooled portfolio against the benchmark by percentage returns. In Figure 3.3, the orange histograms represent the empirical distribution of bond total return and in blue the pooled portfolio performance. The plots regards four semesters, respectively December 2017, December 2018, June and December 2020.[13]

Figure 3.3. Total Return distribution for Benchmark and Portfolio

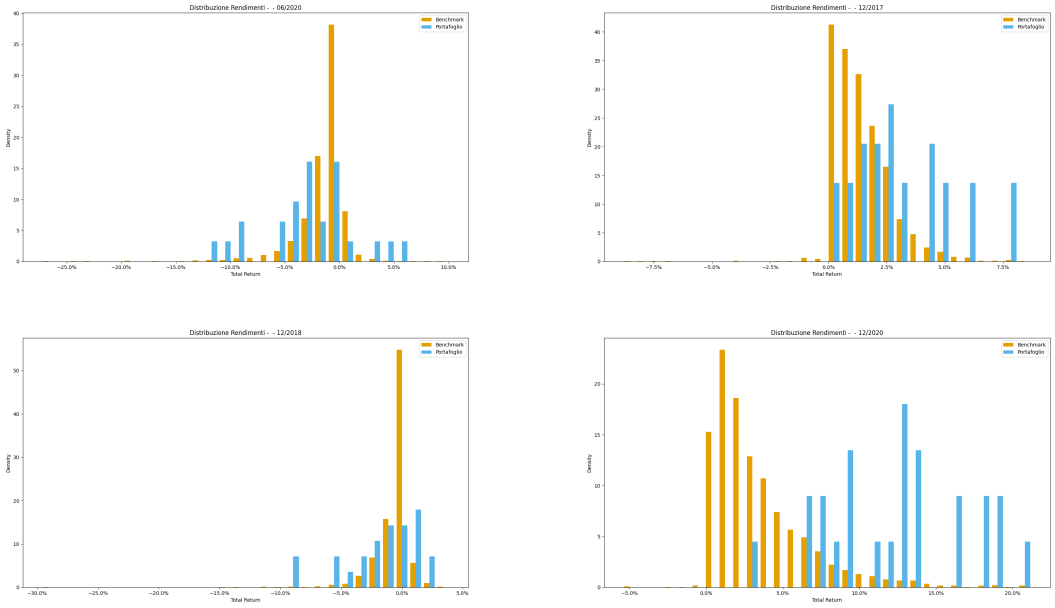


Figure (3.3) presents the histogram of the total return distribution for Benchmark and Pooled Portfolio (Pooled BC) in four noticeable periods during the walk forward excercise from 2016 until 2020. The top left figure shows June 2020, the top right figure December 2017. The bottom left distribution present the evidence for December 2018 and the bottom right for December 2020. In each figure, orange bar represent the benchmark distribution whereas the blue ones the portfolio distribution.

Considering the different periods, we start by analyzing the ones in which the portfolio performed less then the benchmark. In December 2018, the resulting port-

[13]We didn't present the whole sample given that the remaining ones can be associated with the ones plotted, and thus don't retain more information. For seek of completeness, we plotted the two periods in which the portfolio under-performed the most the benchmark and the period of maximum gain.

folio exhibits some important losses that are not offset by the relative high amount of good choices of bond with positive total returns. Is this setting the distribution of the portfolio is more skewed compared to the benchmark. In June 2020, the semester hit by Covid crisis, the distribution is not far from the benchmark and the losses are capped at 10%, with the choice of some sectors that were severely affected by the new normality, such as travels and real estates, and we recall such an event was impossible to forecast from a quantitative based model.
We emphasize the portfolio returns' distribution in the semesters in which it well outperformed the benchmark, namely December 2017 and December 2020. It is possible to see that the algorithm manages to select the bonds that perform the most very accurately. In these instances, the total return distribution presents an asymmetric shape characterized by very limited number of bond recording a negative performance.

The last analysis we carry is to properly estimate how the algorithm selects winner and losers. Figure (3.4) shows the sectorial composition of the benchmark and the portfolio. Again, we plot just four examples of the distribution of the pooled portfolio against the benchmark. It is possible to see that the algorithm tends to replicate the composition of the index with over and under weight in certain sectors, that are not fixed along the time. Benchmark allocation doesn't change rapidly, whereas for the portfolio this doesn't happen. From December 2016 to June 2017, the share of bond invested in Financial sector diminishes from 35.2% to 24.6%, offset by increases in Technology, Transportation and Services. This changes in the portfolio composition represent the extent with which the algorithm expresses its view on each sector performance against the benchmark.

Once we assessed the potential of the procedure, we run a second exercise, by building four portfolios at different percentiles of the forecasted distribution. In such a way we check whether the algorithm correctly specifies the different aspects of

Figure 3.4. Portfolio and Benchmark Distribution by sector

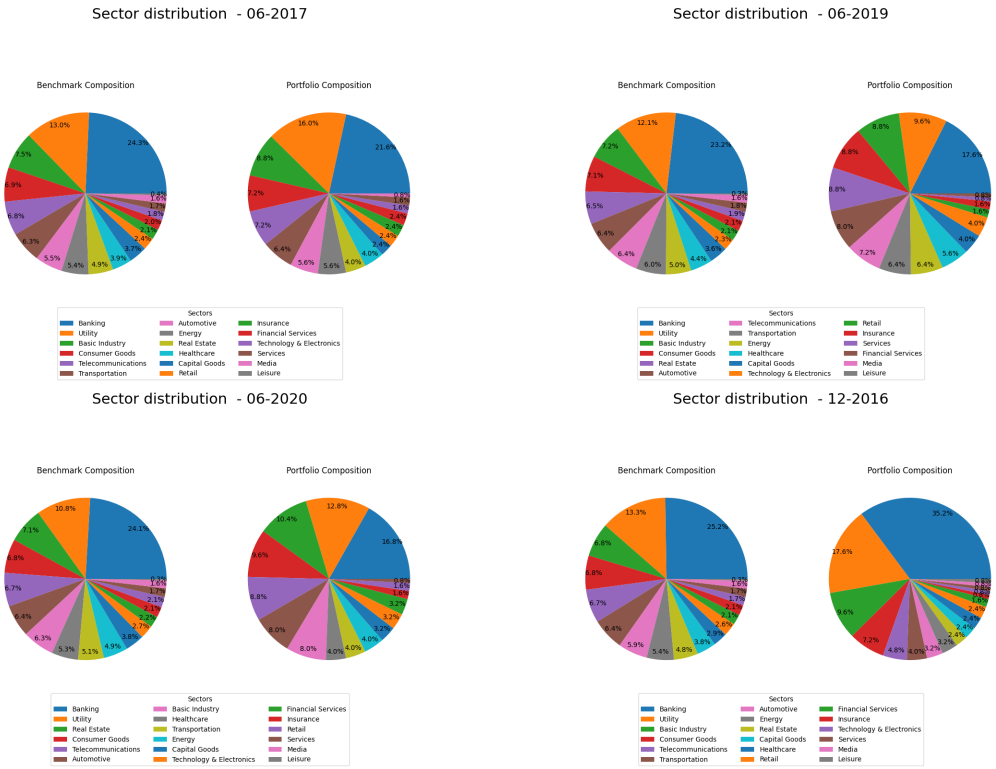


Figure (3.4) shows the pie plot presenting the sectorial distribution of the benchmark and the portfolio in four semesters along the walk forward exercise. The four semesters represent clue moments, namely: June 2017, June 2019, December 2016, June 2020. In each subfigure, the pie on the left represents the benchmark while the pie on the right the portfolio composition.

the distribution. This is a more challenging exercise given that we check the entire distribution to assess whether we are correctly characterizing quintiles. We employ the same walk forward exercise run but we select four portfolios. Q1 is the portfolio of the bond in the first 25% of the forecasted distribution, Q2 is the portfolio of the bonds in the second 25%, Q3 is the third portfolio with bonds between the 50 and 75^{th} percentile and Q4 is the lower 25% of bonds in each semester. Table 3.12 shows the result. In the first columns we check that the number of bonds in each portfolio is similar. We then dive into portfolio performances. Q1 results to be the best portfolio in all semester with exception of December 2018 and June 2020. Also in June 2018 it is not the best performer, but the difference with the top performer

Table 3.12. Walk Forward exercises considering four portfolio

	Number of Bonds				Total Return				% over benchmark				Portfolio Result			
Date	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Dec-16	401	402	402	401	1.3%	0.2%	-0.2%	-0.6%	83%	56%	40%	18%	1.01	1.00	1.00	0.99
Jun-17	437	436	435	436	1.1%	0.2%	0.0%	-0.3%	74%	31%	20%	10%	1.02	1.00	1.00	0.99
Dec-17	471	470	470	470	1.8%	1.5%	1.3%	1.0%	53%	44%	36%	26%	1.04	1.02	1.01	1.00
Jun-18	505	506	505	504	-0.1%	-0.1%	-0.1%	0.1%	65%	66%	67%	75%	1.04	1.02	1.01	1.00
Dec-18	533	534	534	532	-1.1%	-0.7%	-0.3%	-0.2%	52%	67%	79%	87%	1.03	1.01	1.01	1.00
Jun-19	552	551	553	549	5.6%	4.4%	4.2%	4.1%	53%	42%	37%	38%	1.09	1.06	1.05	1.04
Dec-19	564	564	562	561	1.1%	0.2%	0.0%	-0.1%	64%	25%	18%	15%	1.10	1.06	1.05	1.04
Jun-20	524	524	522	522	-2.4%	-1.1%	-0.8%	-0.6%	39%	70%	82%	90%	1.07	1.04	1.04	1.03
Dec-20	477	478	477	475	5.1%	3.0%	2.4%	2.1%	62%	32%	20%	14%	1.13	1.08	1.07	1.06
Avg	496	496	496	494	1.4%	0.9%	0.7%	0.6%	61%	51%	48%	41%				

Table(3.12) shows the walk forward exercises considering four portfolio based on percentile ranking with semester rebalancing. Q1 is the portfolio composed by bonds that the algorithm collocates between the 1^{st} and the 25^{th} percentile. Q2 is portfolio with emissions between the 26^{th} and the 50^{th} percentile. Q3 is the portfolio with emissions between the 51^{st} and the 75^{th} while Q4 is the portfolio of bond from the 75^{th} until the last percentile. The number of bonds represents the number of bonds in each portfolio. Total return is computed equally weighting the bonds in the portfolio. % is computed nominally on the benchmark. Portfolio Results is computed compounding Total Returns without Transaction Cost and is the only measure for which averaging does not add information.

is negligible. In addition, excluding these negative periods, it's possible to note that the returns of the four portfolios are correctly ranked: the algorithm is thus able to distinguish, other than the top performer, the relative order of the other portfolios. When looking at the percentage of bond that over-perform the benchmark, there is the same relationship observed in semester total return, with Q1 well above the others. This indicates that the algorithm allows to asses other then the performance of the single emissions, their correct specification

3.6 Climatic Metrics

We then extend the analysis by considering in the issuer level dataset variables that express the climate risk linked to the issuer, regarding the transitional risk component.

Transition risk variables have been found to affect equity instruments through a risk premium demanded by investors, as classical asset pricing models suggest (Bolton and Kacperczyk (2020)). However, bond investors are intrinsically exposed to downside

risk, so the documented larger hedging cost in the options market for the carbon-intensive firm (Ilhan et al. (2021)), should affect fixed-income investors. In addition, lower returns associated with bonds issued by polluting firms(Duan et al. (2020)), implicitly suggest that investors face the risk of a higher price to bear the firm's credit risk.
However, these exercises assess the current impact of transition risk variables rather than foreseeing corporate bond returns. The proposed approach, disentangling the idiosyncratic change in firm credit riskiness, could overcome the previous limitation. The main reason is that climatic variables affect just the residual components of corporate bond returns, with prevailing effect in some sectors.

We check whether climate risk variables are chosen by the random forest algorithm as a driver for bond overperformance. In this light, a variety of transition risk-related variables can be associated by the non-linear capabilities of the statistical model to a positive or negative variation in the idiosyncratic change in the asset swap. We collect for each firm the Scope 1, Scope 2, Scope 3 Emission as well as their yearly change. These data are however available at yearly frequency whereas the analysis carried out has a semester outlook. We don't restrain or modify the procedure but we analyze whether there is a difference in the results comparing the different semesters. If any information regarding a firm's emission affects prices, it acts especially when the information becomes available. We replicate the procedure described in section (3.5.3) and, rather than the portfolio total return performance, our focus concerns the relative importance of the newly introduced variable with regard to the ones already discussed. Table 3.13 presents impact of adding transitional risk variables into the features set on model M2 presented in section (3.5.2), that include firm-level variables.

In particular, Table (3.13) present the ranking of the transition risk-related variables in the forecasting procedure for changes in $e\Delta_{ASW}$. We present the results for the two semesters distinctly because we want to test whether the release of emission

Table 3.13. Variables Importance for transition risk variables

	Jun		Dec	
	(1)	(2)	(1)	(2)
% Analyzed	79.5%	48.7%	79.5%	50.2%
# Features	20	28	20	28
	Rank			
Scope 1		15		17
Scope 2		19		22
Delta Scope 1		12		19
Delta Scope 2		27		24
Scope 1 Adj		26		25
Scope 2 Adj.		28		26
Delta Scope 1 Adj		20		28
Delta Scope 2 Adj		21		27

Table(3.13) relates to the updated walk forward exercises employing climate transition risk variables. It builds on M2, that considers just equity variables. % Analyzed represent the average number of bond for which the algorithm is able to produce a forecast. Different levels between Jun and December derive from the fact that December has one more observation - December 2017. Feature importance is derived by considering the Gini coefficient, that ranks features according to the impurity reduction at each node employing the Gini Index(Nembrini et al. (2018)

information affects the algorithm's performance.

The first evidence we document regards the percentage of the analyzed bond, that declines with respect to the M2 algorithm. The reason behind the drop resides in the fact that for a given firm, just missing information on any emission variables undermine the possibility for the algorithm to project the idiosyncratic change in ASW.

The number in the columns rank represents the relative importance of each variable in the selection algorithm. A lower number represents more salient information carried by the selected variable. A 1 would mean that the random forest algorithm chooses that variable as a major driver for future changes in $e\Delta_{ASW}$.

The results suggest that the climate variables are still not able to address projection for idiosyncratic change in an asset swap.

3.7 Conclusion

In this work, we derive the risk driver for bond excess return in an exercise of bond selection. Idiosyncratic change in Asset swap, $E\Delta_{ASW}$ is the only variable to project in order to correctly forecast the excess return of a bond with respect to its duration-matched benchmark. To forecast this value, multiple data sources are assessed and pooled. Issue and issuer characteristics, in addition to their evolution over time, can forecast the change in asset swap, especially when analyzed through a random forest algorithm, that exploits non-structured links in the features.

Finally, deriving the excess return from the forecasted idiosyncratic asset swap variation allows building a portfolio that over-performs the benchmark in all but two down-turning periods, where it can manage the losses.

Considering the four portfolios composed of bonds in different percentiles, there is a compelling ability of the algorithm in designating the top and worst performers. Finally, including transitional risk variables, we find that their exploitative power is still very limited when employed in the forecasting exercise.

Chapter 4

Credit Risk Decarbonization: evidence in corporate bond market

4.1 Introduction

Global warming is a certain threat to the environment and the economics and recent evidence highlights its acceleration (Xu et al. (2018)). GreenHouse Gas emissions(GHG) are the main element affecting the earth radiating power(Hsiang and Kopp (2018)) with the main polluting element, carbon emission CO_2, stemming from electricity generation (Stern (2008)). The way to prevent it is by diminishing CO_2 through the transition to a net-zero economy. Nowadays, regulators are surveying diverse ways to curb emissions, from Pigouvian taxation to offset externalities (Baiardi and Menegatti (2011)) to another type of emission regulation (Green (2021)). Transitional risk emerge as regulatory risk (Hjort (2016)) and is still under a process of translation into action. There are however sectors with active regulation, such as the EU ETS cap and trade system employed in Europe for utility firms. Through this mechanism, each year auction mechanisms settle the amount of emission and firms emitting more have to purchase certificates. Even if a reduction in CO_2 emission is considered the first step to tackle as a corporate strategy to mitigate climate change (Kolk and Pinkse (2005)), resulting consequences for firm valuation have not always

coincided (Lee et al. (2015)).

Research has investigated mainly equity instruments, finding investors to value environmental performances(Clarkson et al. (2011)) as well as CO_2 reduction (Bolton and Kacperczyk (2020)). However, equity and bond investors face different payoff structures, so transition risk impacts differently financial instruments (Chen and Silva Gao (2012)). When the analysis regards bonds instruments, green bonds represent a natural framework to build up. A well employed technique is to set up an empirical strategy to match green and conventional bonds to measure the spread with different results. Green bonds exhibit a negative premium compared to the normal one(Zerbib (2019)). When the analysis contemplates just states and other governmental entities, the greenium vanishes as investors account green and conventional bond as substitutes (Lebelle et al. (2020)). Credit risk has been analyzed in the pulp, paper, and chemical industry to test an impact of environmental performance on bond pricing Schneider (2011) and as a directly influencedHöck et al. (2020)

The main question that remains is whether becoming greener pays-off, and it's possible to answer just understanding in which situation it does(Russo and Minto (2012)). This is our central question, by taking advantage of the EU ETS program to understand when the market perceives a premium by greening operations.

Our main contribution is the exploitation of the transition risk horizon in the corporate bond market. We analyze firms in the utility sector that issue a corporate bond with an investment-grade rating and assess whether the duration plays a role when analyzing transition risk. Each year, we split firms into quintiles according to the variation in carbon emission. This analysis enables to discard of major risk drivers for bond performance and retrain carbon emission as the main difference. We analyze for the top and the bottom quintiles the bond split into long and short duration buckets. The objective is a pairwise comparison and we find that the portfolio of green bonds outperforms the brown one, just in the short duration bucket. We then analyze the change in asset swap spread, a risk driver for bond performance. Our conjecture is that transition risk, acting as a regulatory kind of risk, is salient in

the short duration. Controlling for credit risk-related variables, we find overall that change in emission is a credit driver and statistically significant especially for the short-term bond duration. We then control for possible other characterization for transition risk drivers, confirming that change in Scope 1 emission is the key factor in the utility industry.

4.2 Methodology

4.2.1 Transition risk buckets

Transition risk arises from the prospective of a sudden change in the regulatory framework to shift the economy to a net zero target(Grippa et al. (2019)). Industrial organizations are exposed to this kind of risk and the question is through which channel this risk propagates(Battiston et al. (2017)). Firms however are not a standalone in the financial markets and the devaluation of distressed firms' financial instruments could propagate in the economy. The quest is to analyze different financial instrument and understand the specific impact that transition risk poses. When analysing equity, the discussion is the variable of interest, whether the level or the change in CO_2 emission (Bolton and Kacperczyk (2020)). When considering EU ETS regulation, the yearly reduction in the total amount of auctioned CO_2 allowance calls for a correspondent reduction in firm emissions. In this context emission exposure acts as a regulatory kind of risk (Blyth et al. (2007); Cavanagh et al. (1993)) and we test whether the risk arise with the path to diminish CO_2 emission of each firm. For each firm, we indicate with:

$$\Delta_E = \frac{E_t}{E_{t-1}} - 1 \tag{4.1}$$

the yearly percentage change in emission reduction where E_t represents Scope1 emission at time t. The critical point to our argument is the definition of the risk exposure. We do not consider either revenues standardization nor emission because of the characteristic of the EU ETS program that doesn't allow increase in European

energy total consumption. In addition, the percentage of auctioned allowances is almost universal, but still a slight minority and there is a residual percentage of allowances still allocated. A second reason is the fact that we want to isolate the emission reduction effect: if a firm decreases its emission following a crisis, then one would not expect to see a positive portfolio performance nor an improvement in its creditworthiness. It is possible to rank in each year firms according to the Δ_E and select the top and the worst to obtain the two groups to analyze. We indicate the firm in the top quantile in reducing CO_2 emission *green* while with *brown* we indicate the firm in the last quantile.

Equity instrument are the most analyzed given their availability, whereas on bond market salient analysis are carried analysing the differences among green and normal bond. Debt instruments face an higher order of risk component compared to equity, and we address poor results steam from this ultimate reason.

Our analysis focuses on Euro-denominated bond with investment grade rating (BBB or higher) and fixed coupon, which are the components of BofA ML Euro Corporate IG, ER00, bond index tracked by ICE. We thus have all bonds with the same underlying risk-free factor, paired riskiness (no high yield component) and same cash flow structure(no floating rate [1])

To shrink the size and risk components we selected just one sector in which to conduct the analysis.[2] Recalling the aim of the work, that is to analyze the influence of regulatory risk in credit risk, we examine issuers in the utility sector, affected by the greatest source of regulatory risk in Europe that is EU ETS.

This analysis faces also a second issue related to the instrument characteristics. Contrary to equity, each bond has a precise payoff structure in the future, and each expected payment shifts associated risk in different futures outcomes. We split bond in duration buckets according to the regulatory risk horizon. Regulatory risk impacts investment decisions when is enforced with a short period horizon up to 5 years (Gros

[1]Fixed to floater are present in the index provided they are callable within the fixed-rate period and are at least one year from the last call before the date the bond transitions from a fixed to a floating rate security

[2]Accordin to Sector 3 definition provided by ICE

et al. (2016)). If a firm survives the enforcement period, the probability of default lowers with the time, displaying the same shape of the credit risk curve for a firm close to bankruptcy. We aim to test whether this kind of transition risk is priced differently according to the bond payoff structure. We split bonds by duration at the beginning of each year, D_b and setting the threshold between short and long duration at 5 years.

Combining the two separations we made, the first at issuer level between green and brown firms and the second at bond level between long and short duration, we obtain 4 portfolios.

- Green portfolio, Short-duration(GS) : portfolio containing the bond in ER00 issued by one of the top 5 firm in reducing Scope1 CO_2 emission in that year. Duration is between 1.5 and 5.5 years.

- Brown portfolio, Short-duration(BS) : portfolio composed by bond in ER00 issued by one of the worst 5 firm in reducing Scope 1 emission each year. Duration is between 1.5 and 5.5 years.

- Green portfolio, Long-duration(GL) : portfolio composed by bond in ER00 emitted by one of the Top 5 firm in reducing Scope 1 emission in each year, duration between 5.5 and 10.5 years.

- Brown portfolio, Long-duration(BL): portfolio composed by bond in ER00 emitted by one of the Worst 5 firm in reducing Scope 1 emission in each year, duration between 5.5 and 10.5 years.

As anticipated bond portfolios have been divided according to the duration to worst[3] because it is the main driver in price change and thus in bond performance. The alternative was a subdivision by the maturity that would have pose challenging questions for callable bonds. We recall that the use of a fixed rate gives the advantage that the change in prices reflect changes in firm riskiness and is also possible to

[3]Duration to worst takes into account callable bonds by computing the duration for the worst case prepayment option the investor can face.

obtain the ASW for each bond, reflecting the risky component.

4.2.2 Portfolio assessment

Our purpose is a pairwise comparison of green and brown portfolios either in the short-duration bucket or in the long-duration one, detecting the short-duration portfolio as affected by regulatory risk.

Bond Portfolio performances are direct outcomes of change in the level and shape of both risk-free and credit risk interest curve. We point out that the bonds are chosen in a way to maximize similarities of risk factors, to disentangle just possible trade-off coming from decarbonization process. Change in level and shape of risk-free curve affects in an identical way all the bonds in the portfolio, provided that they belong to the same duration bucket. The main driver for the credit risk is change in ASW, a derivative that disentangle the price of credit as a spread over LIBOR.

It is possible to decompose total return for a constant maturity index in the following way, looking at the so-called spread duration Fabozzi and Mann (2012)

$$R = Ytm - D * (\Delta_{RF} + \Delta_{ASW}) \tag{4.2}$$

where Ytm is the yield to maturity of the index, D is the duration, Δ_{RF} and Δ_{ASW} are respectively the change in the risk-free rate and the change in ASW at the corresponding duration of the Index. The most common derivative used to calibrate and measure credit risk is Credit Default Swap (CDS). CDS however represent a minority share (when not completely absent) of institutional portfolio, and such motivation drive us to measure the empirical credit risk decarbonization that can be found in portfolio. From the expansion in Equation (4.2), we are missing higher moments than the second, including convexity but for now, it is fine for two reasons. The first one is that we are not going to fall into the "convexity trap" Gilkeson and Smith (1992) thus having an opposite impact on total return from ASW compared to duration. The second one is that in the first place we want just to analyze the impact of

emission reduction on firm riskiness, and here we are measuring more this effect than the importance of change in ASW on total return performance. Change in risk-free component is common for all the bond in this analysis so we need to disentangle the change in risky components underlying each bond and analyze it. The second goal of this work is to isolate the idiosyncratic change in risk that results for each firm in the green and brown portfolio. To do so, we have to consider the common risk factor still in our portfolio that affects bond, utility sector risk, and net it from each bond ASW. We thus compute the "benchmark" variation in ASW at each year for all duration buckets ($BMK\Delta_{ASW,t,Y}$). This common risk factor, change in ASW, depends on levels and curvature of the ASW curve that is a function of maturity, and hence duration. To be as precise as possible and release the true idiosyncratic change in ASW, we should split the sample in duration bucket the smaller the possible, taking into account the tradeoff consisting of not enough bonds for a good benchmark. The final decision is for a 1-year duration bucket and so it is possible to express the benchmark for change in ASW at time t for bucket duration D as:

$$BMK\Delta_{ASW,t,Y} = \frac{\sum_{i=1}^{n} \Delta_{ASW_b,t} 1_{D_B=Y}}{\sum_{i=1}^{n} 1_{D_B=Y}} \tag{4.3}$$

where 1 represents the indicator function to include for benchmark Y just bonds with this duration. The benchmark can thus be considered as a systematic change in delta ASW and from this, it is possible to recover the idiosyncratic factor by subtracting benchmark in [4.3] from bond ASW change:

$$ID\Delta_{ASW,t,Y} = \Delta_{ASW_b,t,D} - BMK\Delta_{ASW,t,Y=D_B} \tag{4.4}$$

We call this idiosyncratic change because it has been obtained disentangling common risk factors with other bond and is lastly the object of our analysis. We analyze the resulting idiosyncratic Δ_{ASW} at bond level,in a pairwise comparison carried out at two duration buckets of brown and green portfolios.

4.3 Data

In this work we employ data from 31/12/2013 until 31/12/2020. This time-span coincides with phase III of the EU ETS which started in 2013 where the total amount of emissions decrease year by year. We start our analysis from the BofA Merry Linch Euro Bond Investment grade index, ER00, that contains the bond traded in euro with an Investment grade rating and a fixed coupon structure. Bond selected are those with Third level sector equal to 'Utility'. The number of selected emissions varies year by year depending on the firm selected, but in agreement with the quantile selection approach, the total number represents merely the 20% of the bond in this sector. From the index, it is possible to extrapolate the constituent firms from 2013 to 2020 through the issuer parameter associated at each bond. For this firms, we obtained the value of Scope 1 Emission at each year in our analysis from Thomson Reuters and thus determine the Δ_E measure employed in Equation (4.1). To validate our findings, we run the model also employing E/S to match our results with Bolton Bolton and Kacperczyk (2020). Bonds are assigned to each duration bucket according to the duration the bond presents at the beginning of the period. It would be possible to project duration considering only the flow of time but in this regards it is not necessary given that we are not exploit performances. Prices and bond characteristics, including ASW that we employ in the last section, are obtained through the ICE database and compared with Eikon database finding no relevant differences.
Portfolio performances are computed equally weighting bond's total return performance. Hence it is possible to derive portfolio performance since 2013.

Table 4.3 shows the composition of bond database into short and long duration. The share of bond in the short duration bucket is larger compared to the long duration one but increases more proportionally in the last years due to the lower interest rates environment. The split at 5 years duration bucket doesn't derive from an optimization exercise. If it is lower, in the range from 4 to 5 years, results still holds. When it is longer, the long duration bucket becomes non-informative given the small number of

Table 4.1. Bond Emission summary statistics

Panel (A) Index statistics

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
No. Bond	87	91	87	89	91	97	105	67	65	80	75	96	93	99
total return	6.3	0.7	4.2	1.8	-0.2	4.3	1.2	14.2	-0.4	6.6	2.7	-1.2	9.1	5.1
std	2.7	1.5	3.1	2.6	1.0	4.2	3.1	3.1	2.4	2.1	3.9	2.1	3.8	2.6
duration	4.0	4.1	4.0	3.9	4.0	4.1	4.2	6.9	7.1	7.1	7.1	7.1	7.1	7.2
total return B	6.81	0.9	4.8	2.5	-0.4	5.1	1.4	14.5	0.3	6.9	3.8	-1.7	10.1	5.6
total return A	5.4	0.4	3.1	0.5	0.3	2.2	0.8	13.9	-0.7	6.7	1.1	0.1	7.1	3.9

Panel (B) Portfolio characteristics

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
No. Green	6	15	11	7	12	12	21	4	7	4	5	14	9	17
No. Brown	17	10	8	12	7	8	17	4	5	12	5	9	12	9
Duration Green	4.1	3.9	3.8	3.9	3.7	3.9		6.6	7.6	7.3	7.4	6.9	7.4	6.7
Duration Brown	4.2	4.0	3.8	3.6	3.3	3.8	3.4	6.7	7.1	7.4	7.2	7.8	7.3	6.3

Table (4.1) shows the the Index emissions characteristics in the various years. Panel (A) presents the number of bonds (No. Bond) that are analyzed in each duration bucket, total return that represents the yearly total return, equally weighting the components. Std indicates the standard deviation of the total return distribution. Duration is the average of bond duration to worst, computed with respect to worst case scenario dor the investor. Rotal return B represents the average total return with a B grade rating, whereas total return A is the specular measure bond rated A. Panel B presents the characteristics of each portfolio in the different duration bucket.

bonds.

4.4 Portfolio Performances

To obtain portfolio performances we evaluate from time $t-1$ to time t the emission pattern of a firm. At time t we compute issuer's Δ_E to select green and brown firm. We extract the bond issued by those firm and, among both green and brown portfolio, we split in short and long duration to form the four portfolios, GS, BS, GL, B. and evaluate their total return performances(TR) [4] from time t until time $t+1$. Table 4.3 shows the result for the 4 portfolios.

The first fact one can note is that while in the short-duration the green portfolio outperforms the brown one in all years but 2014, in the long-duration portfolio the

[4]Total Return takes into account both change in price and dividends

Table 4.2. Issuer Emission Summary statistics

	2014	2015	2016	2017	2018	2019	2020
No. Firms	28	26	29	27	30	34	32
E	3.6E+07	4.5E+07	4.2E+07	2.5E+07	2.7E+07	2.4E+07	2.3E+07
σ_E	5.4E+07	6.0E+07	6.1E+07	4.3E+07	5.0E+07	4.8E+07	4.7E+07
Δ_E	-1.0%	-5.7%	-4.5%	-9.5%	-8.7%	-6.4%	-7.5%
$\sigma_{\Delta E}$	27.1%	12.1%	23.7%	25.4%	22.9%	22.1%	20.2%

Table (4.2) shows the Emission related statistics for the firms composing the database. No. Firms is the total number of firms in the database each yeat, E represents the estimated tonnes of CO_2 averaged emitted by firms, equally weighting the components. Δ_E represent the average yearly perccentage change in E and $\delta_{\Delta E}$ is the yearly standard deviation.

same doesn't occur. In the long duration indeed, we observe an alternation in the performance between the green and brown firms. For the same firm then, there is a comparative advantage in investing in the short-duration bond in opposition to long-duration one. The only concern regards any possible missing factor that could explain the discrepancy in portfolio performances. We investigate the main drivers that impact on portfolio performances to check whether the division between green and brown firm has influenced this result.

The first effect we check is the duration, that act as amplifier in the bond total return. In the short-duration bucket the duration of the two portfolios ranges between 4.1 and 3.3 years and is there are no considerable differences between the two. There is only one year, 2018, which displays a substantial difference between the two measures, with the green portfolio characterized by a duration of 0.47 years longer than the brown portfolio. The difference in the performance however is remarkable since the worst portfolio has a negative performance while the top performer has a positive performance. The difference in performance is not caused by the longer duration because its property is to amplify changes in yield to maturity, while preserving the sign.

Considering the long-duration bucket instead, the duration extends from 6.5 years to 7.8 years. Also in this context, in 2018, the two portfolios experience a difference in duration: 6.9 years for the green firm against 7.7 years for the worst ones, almost

Table 4.3. Portfolio performance statistics

Panel (A) Portfolio total return, percentage point

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.0	**1.3**	**3.7**	**2.2**	**0.1**	**3.7**	**1.3**	**13.3**	0.7	**7.0**	1.3	**-0.3**	7.7	3.2
Brown	**7.3**	1.0	2.5	2.0	-0.2	2.5	1.2	12.3	**1.1**	5.6	**5.6**	-1.4	**7.9**	**3.7**
G-B	-1.3	0.3	1.2	0.2	0.3	1.2	0.1	1.0	-0.4	1.4	-4.2	1.1	-0.2	-0.5

Panel (B) Portfolio Duration

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	4.1	3.9	3.8	3.9	3.7	3.9		6.6	7.6	7.3	7.4	6.9	7.4	6.7
Brown	4.2	4.0	3.8	3.6	3.3	3.8	3.4	6.7	7.1	7.4	7.2	7.8	7.3	6.3

Panel (C) Benchmark total return, percentage point

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.8	0.9	4.8	1.6	-0.4	3.4	1.2	14.5	0.3	6.9	1.1	-1.8	8.4	4.8
Brown	6.8	0.9	4.8	2.5	-0.3	3.7	1.2	14.5	0.3	6.8	3.8	-1.3	8.6	4.8

Table (4.3) shows the results for the portfolio exercise described in section (4.4) . The left panel shows the short duration bucket portfolio while in the right the long duration bucket. Each section of the table present the statistics for the green and the brown portfolio, to quickly asses a pairwise comparison. The first evidence regards portfolios performance, computed as total return. The second part comprises portfolio duration. The final section shows the benchmark total return for each portfolio, computed according to equation (4.5). Portfolio performance in bold shows the top performer for each duration bucket.

one year of difference in duration. In this case, the green portfolio outperforms the brown one, but part of this difference has to be attributed to the lower duration a longer duration implies a magnified effect of changes in yield on bond's performance and, given the negative return caused by an increase in the yield to maturity, a lower return.

A second stage of analysis is to compare these results with the portfolio benchmark: the results suggest that only in the short-duration framework green portfolios outperforms brown ones, but it is still possible the benchmark to perform even better. In this case, a positive performance of the first quintile would be lower than the benchmark, a particular situation that prevent a positive outcome when investing in the green portfolio. Through the previous procedures, we isolate a sector and duration bucket, so a possible risk driver can be represented by credit rating (Cochrane

and Piazzesi (2005); Pogue and Soldofsky (1969)). Various market conditions reward differently firms credit risk. As shown in Table (4.1), years characterized by positive market performance display a higher returns bond with a higher credit risk. The opposite happen in downtuning periods, when A-rated bond are purchased as safe asset.

To conduct a rigorous analysis, we match the portfolios against its rating-based benchmark. We split the bonds in two subcategories according to their ratings: Rating A (AAA - A) and Rating BBB.

For each rating class, we compute the yearly average returns in the long and the short duration bucket. For each portfolio, we compute weighted average of the benchmark return, R_b, as:

$$R_b = R_A * w_A + R_{BBB} * w_{BBB} \tag{4.5}$$

where R_A and R_{BBB} are respectively the rating based portfolio performances for the bond with A and BBB benchmark. The weights, w_A and w_{BBB}, represents the percentage of bonds in the portfolio with that rating.

At each time there are four different benchmarks, each one related to a single portfolio. The Panel (C) in Table (4.3) shows the total return for the rating based benchmark in the long and in the short duration bucket for the green and the brown portfolios. Two benchmarks coincides in the case the proportion of bonds in the two rating class coincides.

The aim of the analysis is to disentangle two possible drawbacks in the analysis. Firstly, a positive overperformance for the green portfolio in the short duration bucket can derive from a particular choice of bonds with a certain rating. Secondly, the overperformance of the green bonds just against the brown ones would imply that presence of bonds that are able to disentangle such an overperformance.

We analyze the last section of Table (4.3) starting from the short duration bucket. The benchmark total return is almost equal in all the years with 2017 as the only exception, where the brown portfolio overperformed the green one by 0.9%. It is

possible to note also a minor difference in 2019 again in favor of the brown portfolio. This evidence suggests that the rating composition of the two portfolio is very similar. In particular, analyzing 2017, the brown portfolio is composed mostly by BBB rated bond whereas the green one is more balanced. The BBB rating benchmark, in addition, obtains a higher performance (2.5% versus 1.6% for A) but the return of the green portfolio(2.2%) is still higher then the brown one (2.0%). For the short-duration bucket the green overperformance does not derive from an unbalanced rating selection. Finally, comparing the green portfolio performance against the benchmark, there is evidence of a better performance in all years but the 2014 and 2016. It is possible to exclude also the second possible source of noise.

Regarding the long-duration bucket, the rating based duration buckets exhibits similar values, with also in this case, as it the short duration bucket, 2017 as main difference. In this case however, the brown portfolio well overperforms the benchmark(5.6% versus 3.8%) while the green portfolio obtains a mere 1.3%. In addition. in 2019 and 2020, the green portfolio well underperforms the rating based benchmark.

This finding is in line with recent findings in Australia Qian et al. (2020), where policy changes resulted in higher financial performances for firms with better carbon performances. Policy risk is mainly priced in short term portfolios.

By selecting firms according to their reduction level in CO_2 and splitting the bonds between a short-duration (2-5 years) and long-duration(6-10 years) bucket, we find an overperformance just in the short-duration in all years but in 2014, the first in which phase III for EU ETS was adopted. In the long duration settings, there is no evidence of a best practice. When controlling for possible factors that could mislead our conclusions, from duration to rating composition, there is no evidence of a particular driver for green overperformance. Our assumption is that investors relatively reward green credit risk against brown. In the following section we analyze this implication.

4.5 Variation in Credit Risk

4.5.1 Portfolio Level

It is possible to analyze Idiosyncratic change in asset swap spread, $ID\Delta_{ASW}$, from to two different perspective. When we consider the performance, a positive $ID\Delta_{ASW}$ indicates a higher increase or a lower decrease of bond change in ASW compared to its benchmark, resulting thus in lower bond performance, ceteris paribus. Inversely, it indicates a lower increase or higher decrease that causes a higher bond performance compared to its benchmark. From a risk perspective, recalling that the change in ASW describes the risk attached to that bond, a positive $ID\Delta_{ASW}$ indicates that a firm is evolving in a riskier way than similar firms. A negative value instead represent a bond that is evolving in a less risky way than its similar ones.
We are thus interested in analyzing how $ID\Delta_{ASW}$ behaves according to two different representations:

- changes grouped by portfolios and duration, to compare the changes between portfolios in the same duration.
- expand the research and analyze the change taking into account emission reduction, this could result in determining a risk factor that affects risk.

We proceed with the first task, In 4.1 we compare the boxplot of Idiosyncratic ASW change both in short duration portfolio and long duration portfolio between Top-firm and Worst firm.

Analyzing the situation in the short duration bucket, it is possible to note that the median(Orange line) of Worst firm portfolio is above the box of Top-firm box. This indicates that it is likely for the worst-firms to have a variation in $ID\Delta_{ASW}$ higher than the one of Top-firms, meaning a lower decrease or a greater increase in riskiness compared to Top Firm.
In the Long-duration bucket instead, this effect doesn't show. The two boxplot exhibit similar characteristic and no difference appear to be. The relative riskiness

Figure 4.1. Boxplot of Idiosyncratic change in ASW.

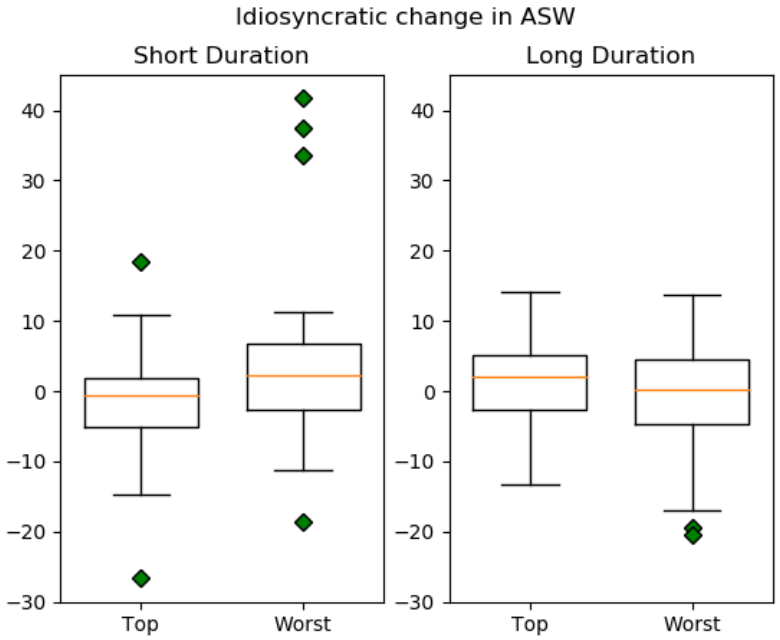


Figure (4.1) presents two subfigures with boxplots for the Idiosyncratic change in Asset Swap Spread (ASW). The left figure plots the short duration bucket framework, comprising bond with a duration lower than 5 years. The figure on the right plots the long duration bucket, that are bond with a duration longer than 5 years. Top represent the green firm an worst the brown one. The box is constructed considering the 25^{th} and 75^{th} percentile

is preserved. These results are in line with what obtained in the previous section: portfolio performance of green portfolio outperforms brown one just in the short duration bucket. For the second task, we collect for the bond in a portfolio, the year change in CO_2 emissions, Δ_E, and the $ID\Delta_{ASW}$. We try to exploit a different relationship in how the change in emission and the change in ASW behave according to bond duration. The model we employ is an OLS that links change in emissions to $ID\Delta_{ASW}$ as follow:

$$ID\Delta_{ASW} = \alpha + \gamma\Delta_E + \epsilon \tag{4.6}$$

We group in two samples, short and long duration bonds to capture the different γ among the two groups. Following the first section, we would expect a positive

Table 4.4. Estimation result

	Short Duration	Long Duration
α	1.02	-0.02
	(0.87)	(0.84)
γ	4.94**	-0.51
	(2.59)	(2.17)
No Observations	144	97
R-Squared	0.073	0.05

T-stats reported in parentheses
***1% significance,**5% significance, *10% significance,

Table (4.4). Sample period is 2014–2020. It shows estimation result for the equation (4.6). Short duration columns runs the regression on the bonds characterized by duration lower than 5. In the Long Duration columns we run the regression on the bond with a duration that is higher then 5 years. The Asset Swap are considered as basis point.

influence of variation in emission on the variation in idiosyncratic ASW change. The smaller the reduction in emission (ore even an in increase) would result in a higher ID movement, meaning an increase of ASW comparative to comparable bond, hence a more increased riskiness.

One can observe that γ for a short duration is greater than 0 at 90% confidence level while γ for a long duration is smaller than 0. Both models have an R-squared score very low, suggesting that alone this variation in spread doesn't explain completely idiosyncratic movement in ASW. It is possible to observe that in the short duration portfolio there is a positive but statistically not significant α coefficient, this means a slightly positive on average idiosyncratic change with respect to long-duration setting. This doesn't alter the previous result: the main point here is the different γ value found in the short and the long-duration buckets. We find the fact of different γ parameters to be very interesting for the development of the research: emission reduction risk, being subject to regulatory risk, concerns more short-duration portfolios rather than long-duration ones. This result could reconcile the view of Byth and Yang on political risk Yang et al. (2008); Blyth et al. (2007) with the one analyzing financial performances Lewandowski (2017).

4.5.2 Aggregate Level

Considering just this small sets of bonds has both positive and negative sides. We show to be able to disentangle portfolio overperformances and change in riskiness but on the other side we were not able to disentangle the effect of CO_2 reduction effects on the entire set. In this section we enlarge the database and look whether there are effects also at an aggregate level. In particular, given the panel data characteristics of the database we could take advantage by considering the time effects adding appropriate controls variables at equity levels. The relationship we estimate is then

$$\Delta_{ASW_i,t} = \beta_e * \Delta_{E_{t-1}} + \beta_c * C + \gamma_t + \epsilon \tag{4.7}$$

where as C we include the DEBT level that measure the amount of Debt the firm is exposed to and the Interest Coverage Ratio (ICR) that returns the percentage of reimbursement to be done covered by cash collateral. We include fixed effects at time level[5] and establish the link through the Panel OLS technique.

As in the previous section, we face a different situation in the long and in the short portfolio duration bucket. After controlling that the coefficient related to debt and ICR are similar and significant[6], it's possible to see that the positive relationship in the short term between changes in Delta Emission and change in ASW fades away in the long run. Again, an increase in emission is linked to an increase in the riskiness of the issuer especially in the short period environment.

[5]considering also fixed effect at the entity level results in overfitting given the sparsity of the data, the average number of observation for each bond is of 4.5 with some with just one or two observations.

[6]An increase in the ICR implies a decrease in the firm credit riskiness, due to the more cash that are available to repay the debt

Table 4.5. Variation in asset swap at duration buckets

	total	long duration	short duration
α	-10.170***	-8.7690**	-11.429***
	(2.0924)	(3.5654)	(2.8788)
Δ_E	41.751***	32.373	49.581***
	(12.577)	(24.470)	(18.498)
Δ_{ICR}	-3.7851**	-7.2017*	-2.9404
	(1.7236)	(3.9861)	(2.2358)
DEBT	11.849	-5.1682	20.386
	(9.4127)	(21.995)	(13.722)
Estimator	PanelOLS	PanelOLS	PanelOLS
No. Observations	872	357	513
Cov. Est.	Robust	Robust	Robust
R-squared	0.1475	0.0748	0.1568

T-stats reported in parentheses

***1% significance,**5% significance, *10% significance,

Table (4.5). Sample period is 2014–2020. It presents the estimation for $\Delta_{ASW_i,t} = \alpha + \beta_e * \Delta_{E_{t-1}} + \beta_c * C + \gamma_t + \epsilon$. Column "Total" presents the result considering the entire dataset. The columns Long Duration and Short duration represent respectively the bucket defined in section 4.2.1 with the former encompassing bond with a duration lower than 5 years, the latter higher. Asset Swap Spread is measured as basis point.

4.6 Robustness

4.6.1 Different transition measures

Transition risk is associated to the circumstance in which the firm's operations and assets become stranded when regulations to coerce CO_2 emissions take place. Which firm characteristic sharpen this exposure has extensively been debated. ESG scores present a section for environmental firm performance that takes into account CO_2 emissions. This could be an alternative green driver. We employ the logarithm Scope 1 Emission to define an alternative

$$Te_{i,t} = log(E_{i,t}) \tag{4.8}$$

Then, we employ also a standardized measure of CO_2 emissions, $E_{i,t}$ over revenues, $R_{i,t}$, as

$$Ts_{i,t} = \frac{E_{i,t}}{R_{i,t}} \tag{4.9}$$

The persistence of risk measure allows the investors to forecast the riskiness of a firm , and is a key also considering carbon transition risk(Duan et al. (2020)). We investigate the transition matrix of different transition risk specification. For each transition risk definition, we split the firms in our sample in quartile and compute the transition probability each year and average the results. Table (4.6) shows the result. Starting from Te, the transition risk defined only on Scope1 level, we note that the diagonal entries are well above 75% and are similar to the credit risk transition matrix. Looking at Te instead the values are slightly smaller, similar to the analysis by Duan et al. (2020) that employs a larger data-sample. Finally, the transition matrix associated with change in Scope 1 emission is even more disperse but in a consistent way, with entries in the diagonal and in the closer section above 25%. We don't address this as a drawback of $T\Delta e$, already Bolton (Bolton and Kacperczyk (2020)) showed that change in Scope1 emission matters more then CO_2 emission itself.

Table 4.6. Average Transition matrix

Panel(A): Te

	1	2	3	4
1	87.4%	12.6%	0.0%	0.0%
2	10.3%	75.9%	13.8%	0.0%
3	0.0%	7.0%	73.6%	19.4%
4	0.0%	3.0%	5.6%	91.4%

Panel(B): Ts

	1	2	3	4
1	76.8%	23.2%	0.0%	0.0%
2	16.7%	59.9%	16.4%	7.0%
3	0.0%	28.5%	57.7%	13.8%
4	3.2%	0.0%	17.7%	79.0%

Panel(C): $T\Delta e$

	1	2	3	4
1	38.3%	27.5%	17.5%	16.7%
2	35.2%	25.9%	25.0%	13.9%
3	30.5%	26.7%	23.8%	19.0%
4	4.6%	17.6%	48.1%	29.6%

Table (4.6). The sample period is 2014-2020. The tables show the average transition matrix of different transition risk measure employed. Transition is computed for issuer that have at least one issued bond in the Investment Grade perimeter of BoFA EG00. Each transition matrix has in the rows quartile at the beginning of a year and in the columns the quartile in the following period.

4.6.2 Including Scope2 and Scope3 emissions

Investors look at the emission profile of a firm but, as pointed out, a firm is not responsible just for Scope 1 emission. Scope1 emissions represent just the direct emissions coming from the operations a firm owns or runs. Scope 2 emission encompass also indirect emissions required to run the operations, such as energy. Finally, Scope 3 emissions include all the indirect emissions linked to the firm's final production, such as transportation of goods. Scope 2 and Scope 3 emissions constitute a challenging exercise concerning their measurement, considering that in certain industries they can constitute 75% of total CO_2 emissions (Downie Downie and Stubbs (2013)), more and more attention is being piad to assess and manage them (Hertwich and Wood (2018)). In utility sector there is already a large evidence that Scope 1 emissions are the most important ones (Lewandowski (2017)). With regard to the database employed, Figure (4.2) shows the relationship between Scope 1 and Scope 2 emissions. The amount of average S1 emissions is in the order of 8 to 10 times the amount of S2 emission.

Considering the perspective of the analysis, what matters is the relative positioning of firms according to the emission level. We test the pairwise confusion matrix between Scope 1, Scope 2 and Scope 3 emissions, with results in Table [4.7].

The results are quite different considering each pair. Looking at S1 and S2, the first and second quartile of S2 are well replicated by S1, with respectively 47.4% and 35% of issuers in the same bucket. The third and fourth quartile however are more disperse, 13.3% of the firms in the bottom 25% in S1 are in the top 25 percentile for S2. Analyzing the dispersion between S1 and S3 or between S3 and S2, we find an higher concentration in the diagonal.

4.6.3 Replication

Considering these possible variations, we run the same exercise in Section 4.4 and 4.5 considering the different specifications regarding the transition risk factors as well as the type of emissions. Table (4.8) presents the results of the different specifications

Table 4.7. Confusion Matrix between Scope1, Scope2 and Scope3 Emission

Panel(A): S2 vs S1

	1	2	3	4
1	47.4%	35.5%	7.7%	9.5%
2	41.8%	34.9%	14.0%	9.3%
3	21.5%	19.9%	37.7%	20.9%
4	13.3%	4.2%	33.2%	49.4%

Panel(B): S3 vs S1

	1	2	3	4
1	70.1%	16.8%	6.1%	7.0%
2	27.4%	24.3%	29.5%	18.7%
3	6.4%	20.4%	47.1%	26.1%
4	3.4%	27.6%	20.7%	48.3%

Panel(C): S2 vs S3

	1	2	3	4
1	52.7%	12.1%	12.2%	23.1%
2	37.6%	32.3%	30.1%	0.0%
3	4.0%	38.8%	25.3%	31.9%
4	0.0%	7.4%	33.5%	59.1%

Table (4.7). The sample period is 2014–2020. It shows the average confusion matrix between the different scope emission ranking in the quartiles employing yearly data. Panel (A) shows the average confusion matrix between Scope 2 emissions (rows) and Scope 1 emissions(columns). Panel (B) shows in the rows Scope 3 emissions while in the columns Scope 1 emission. Panel (C) displays in the rows Scope 2 and in the columns Scope 3.

Figure 4.2. Average Scope 1 and Scoper 2 emission by year

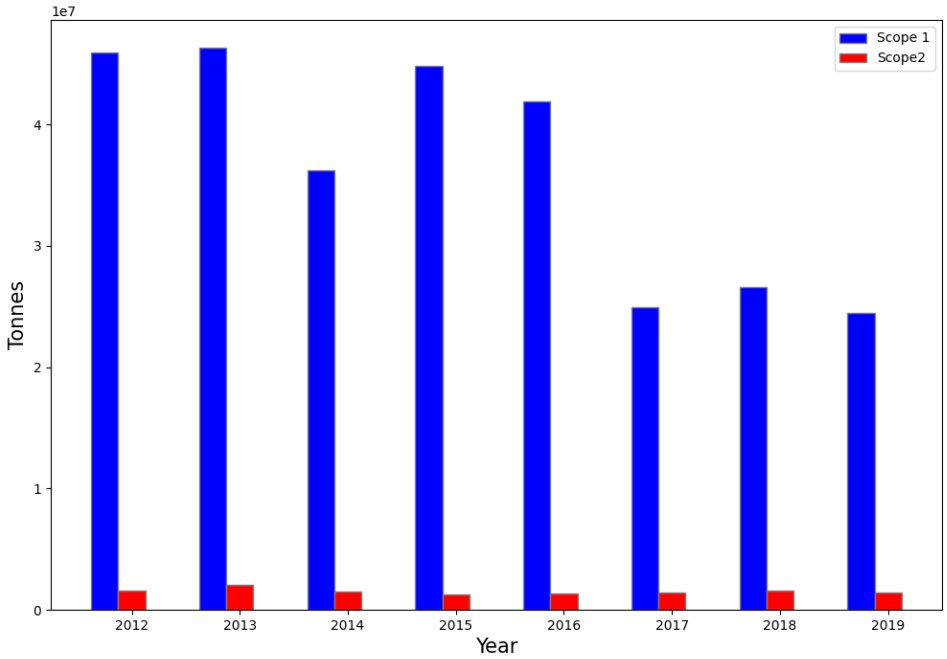


Figure (4.2) shows the average amount of Scope 1, blue bars, and Scope 2, red bars, Emissions in each year. It is computed among the firms in the utility sector that issued at least one bond with investment grade rating in the index BofA EG00.

with a comparison with the portfolio benchmark, Panel(E).

Starting by considering Panel(A) and Panel(B), that take into account the alternative risk metrics such as the total emissions amount, Te, and the standardized emission by sales, Ts, there is no evidence of an portfolio clearly overperforming either in the long and in the short run. When considering Te ranking algorithm in the short duration bucket, the number of time green portfolio overperforms the brown is higher than the number of time it doesn't, and in these situations the under-performance is low. The long short strategy results in a negligible total return. When considering Ts, in the short duration bucket, the green portfolio obtains a positive spread just in the first three years, with a negative performance in the remaining ones. This results are in line with the finding by Bolton and Kacperczyk (2020) that document a risk premium just when considering the change in Scope emission rather then the levels.

Panel(C) and Panel(D) finally show the results employing different Scope Emissions, namely Scope2 in Panel (C) and Scope 3 for panel (D). Again, the results indicate a less persistent overperformance by the scope 2 emission levels. It is possible to explain this fact looking at the sector of analysis and the relation between direct and indirect emission. As shown in Figure (4.2), the amount of indirect emission is considerably lower compared to the direct ones. A slightly change in this component then produce a smaller effect compared to S1. In addition, the EU ETS program acts on direct emissions, and this underlies that a regulatory type of risk acts upon the variables that carries this risk.

Table 4.8. Summary statistics for Portfolio performance

Panel (A) Total emissions transition risk: T_e

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.2	1.1	4.2	1.4	-0.5	3.5	1.2	12.8	0.4	6.8	1.4	-0.1	8.5	4.1
Brown	6.5	0.9	4.3	2.3	-0.2	3.1	1.5	14.6	0.1	6.4	4.2	-1.7	8.6	3.8
G-B	-0.3	0.2	-0.1	-0.9	-0.3	0.4	-0.3	-1.8	0.3	0.4	-2.8	1.6	-0.1	0.3

Panel (B) Sales standardized emission transition risk: Ts

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	7.1	0.9	5.3	1.7	-0.2	3.3	1.4	14.4	0.1	6.9	1.8	-2.1	9.1	3.8
Brown	6.3	0.8	4.8	2.8	0	3.9	1.1	12.2	0.1	6.7	2.9	-0.9	8.2	4.5
G-B	0.8	0.1	0.5	-0.9	-0.2	-0.6	0.3	2.2	0	0.2	-1.1	-1.2	0.9	-0.7

Panel (C): Scope 2 transition risk: $S2$

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.9	1	4.5	1.2	-0.3	3.6	1.3	13.1	1.2	7.5	0.8	-2.1	7.5	4.2
Brown	6.5	0.8	5.2	2.5	-0.3	2.4	1.2	15.1	0.8	6.9	3.8	-1.3	8.5	4.9
G-B	0.4	0.2	-0.7	-1.3	0	1.2	0.1	-2	0.4	0.6	-3	-0.8	-1	-0.7

Panel (D) Scope 3 emission transition risk $S3$

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.2	0.8	3.5	1.8	-0.1	3.6	1.4	14.8	-0.1	6.7	1.8	-1.5	8.1	3.2
Brown	7.4	1.2	2.8	2.1	-0.2	2.9	1.1	11.9	0.1	6.5	3.1	-1.1	7.1	5.1
G-B	-1.2	-0.4	0.7	-0.3	0.1	0.7	0.3	2.9	-0.2	0.2	-1.3	-0.4	1	-1.9

Panel (E) Benchmark

	Short Duration							Long Duration						
	2014	2015	2016	2017	2018	2019	2020	2014	2015	2016	2017	2018	2019	2020
Green	6.8	0.9	4.8	1.6	-0.4	3.4	1.2	14.5	0.3	6.9	1.1	-1.8	8.4	4.8
Brown	6.8	0.9	4.8	2.5	-0.3	3.7	1.2	14.5	0.3	6.8	3.8	-1.3	8.6	4.8

Table (4.8) shows portfolio performance employing different transition risk measures. Panel (A) presents the exercise with total carbon emissions, T_e equation(4.8), as risk measure. Panel (B) shows the results with the standardized carbon emissions, T_s, as defined by equation (4.9). In Panel (C) portfolios are discriminated through Scope2 emissions, while in Panel (D) through Scope3 emissions.

4.7 Conclusions

In this paper, we analyze when it pays off to become green for utility companies, expressed as the time-window horizon in the fixed income market. EU ETS program forces utility companies either to yearly reduce their carbon footprint or to buy

emissions certificates. In this light, firms face a trade-off between higher investment to produce clean energy or undergoing the cost to acquire the certificates. Fixed income instruments, whose payoff exhibit a precise term structure in the payments, are better suited to disentangle a risk horizon concerning equity instruments.

We show that by sorting the firms according to the change in scope1 emission, a long-short strategy is remunerative just when considering bonds with a duration lower than five years. In addition, we analyze the credit risk driver for bond returns in the different portfolios and determine that the positive performance associated with the green portfolios in the short duration is determined by a comparable reduction in the credit risk component. Finally, we link the credit risk performance to the change in Scope emissions and, excluding control variables, we find a positive and statistically significant coefficient suggesting that change in Scope 1 emissions reduces credit risk for borrowers just in the short part of the credit term structure cruve. In this light, the so-called transition risk behaves exactly as a regulatory fashioned kind of risk, with severe impact forecasted when regulation is unexpectedly enforced.

Appendix A

Alternative temperature aggregating methods

In this appendix, we discuss in detail the aggregating methods we develop in Section (2.3.4) to aggregate temperature at the state and country level.

A.1 State aggregation

The first digression regards the state aggregating method. We employ this index to match two analyses. Firstly, from a financial perspective, we employ energy consumption and the cross-section of equity returns considering the firm's headquarter state. The objective is to replicate at the state level the indicator that represents the deviation from average temperature variability, $TD\text{-}VAR$ and allows the comparison with the state daily temperature anomaly, TD. For this purpose, we collect data at the grid level within U.S. space and assign each grid to a particular state.
We then define the state s temperature at day d as

$$T_{s,d} = \sum_{i=1}^{N_s} w_i * T_{i,d} \tag{A.1}$$

that is a weighted average of the temperature in the N_s grids within state s. Implementing this procedure we obtain a coherent measure at the state level with the one defined at the city level. Then, it is possible to proceed as in Equation (2.4) at the city level obtaining the day-to-day temperature anomaly, the day-to-day variability, and the deviation from average variability at the state level. We implement $w_i = \frac{1}{N_s}$ to obtain an equally-weighted state temperature level index. Alternative to the equi-weighting algorithm, it is possible to aggregate based are GDP or population weighted, that will be discussed in more details in the following section.

Since we are dealing with temperature variability, aggregating an index or its variability could result in abnormal differences when the index components exhibits different behaviour[1]. Alternatively to Equation (A.1), there is the chance to directly aggregate state $TD\text{-}VAR$:

$$TD\text{-}VAR_{s,d} = \sum_{i=1}^{N_s} w_i * TD\text{-}VAR_{i,d} \tag{A.2}$$

In this way, the state-level index would represent the average deviation from historical temperature variability. Castellano et al. (2020) shows that even for temperature index aggregation, under suitable conditions, with the main being that T_s follows an AR(p) process, carefully selecting $\sum x_i^2 = 1$ in (A.1) the resulting index has the risk equal to weighted average of the single indexes risk. It is feasible to compare average risk coming from A.2 and A.1.

Another reason that allows us to aggregate temperature and not variability deviation is the number of cell grid within each state. We test possible differences employing city-level data at airport station[2] as temperature data and tested the differences coming from the different aggregating procedures.

Comparing the state deviation variability coming from the aggregate temperature levels and the one coming from the aggregate deviation volatility itself, we find an overall correlation of 95% on the level. Some states exhibit lower values, such as Cal-

[1] $\sigma(\sum w_i * x_i) \neq \sum w_i * \sigma(x_i)$

[2] We enlarge the data set by Diebold and Rudebusch Diebold and Rudebusch (2022) considering all the airports within the U.S. From the NOAA, 330 airports have active data covering 1960 to 2017

ifornia and Texas(but still close to 90%), and are the ones that present more airports in the data set.

We now analyze some characteristics and main differences between state monthly deviation in variability and the monthly temperature deviation. Figure A1 shows the cross-sectional difference for a particular month, September 2005. In the upper section, we plot the monthly state-level temperature deviation (TD) whereas in the lower panel we display the deviation in variability(TD-VAR).

It is possible to note that also at state levels the two measures exhibit different patterns. There are states, such as Texas, that show a high level in temperature deviation as well as a high deviation in variability. States very close to Texas, such as New Mexico and Arizona, display a higher level of deviation in variability (TD-VAR) but, concerning temperature deviation, the levels are among the smallest in the U.S.

A.2 U.S. deviation in variability factor

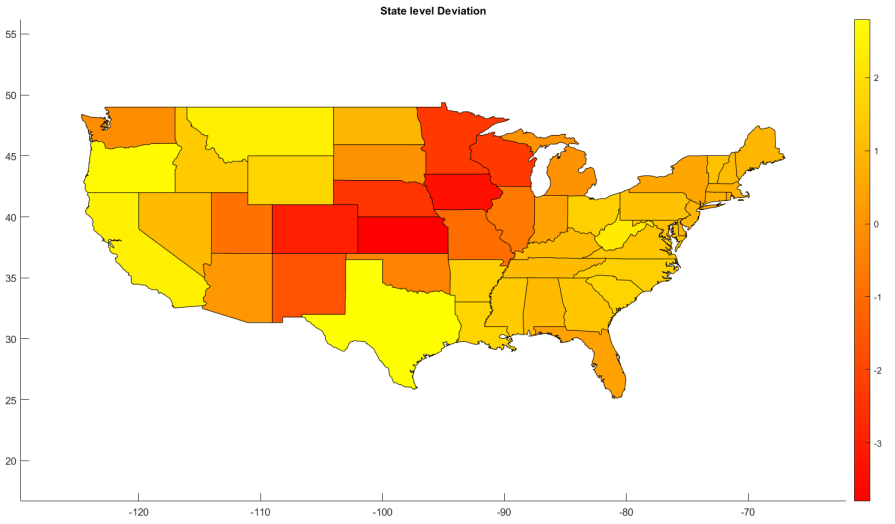
The next step we take is to derive a U.S.-wide TD–VAR factor, that could easily be seen as an aggregate level of the state volatility. Such a method would release the geographical characterization and give the possibility to match against indexes that are not geographically tightened but refers to the U.S. In this context, the aim is to disentangle U.S. TD against U.S. TD–VAR and hence, considering the two possible approaches employed at state level, A.1 and A.2 and that each state can be seen as a single cell, at this stage we build U.S.-wide TD and TD–VAR with A.2, defining

$$US_{TD,m} = \sum_{i=1}^{N_s} w_i TD_{i,m} \tag{A.3}$$

$$US_{TD\text{-}VAR,m} = \sum_{i=1}^{N_s} w_i TD\text{-}VAR_{i,m} \tag{A.4}$$

Appendix Figure A1. State level aggregation for TD and TD–VAR

(A) State level TD



(B) State level TD-VAR

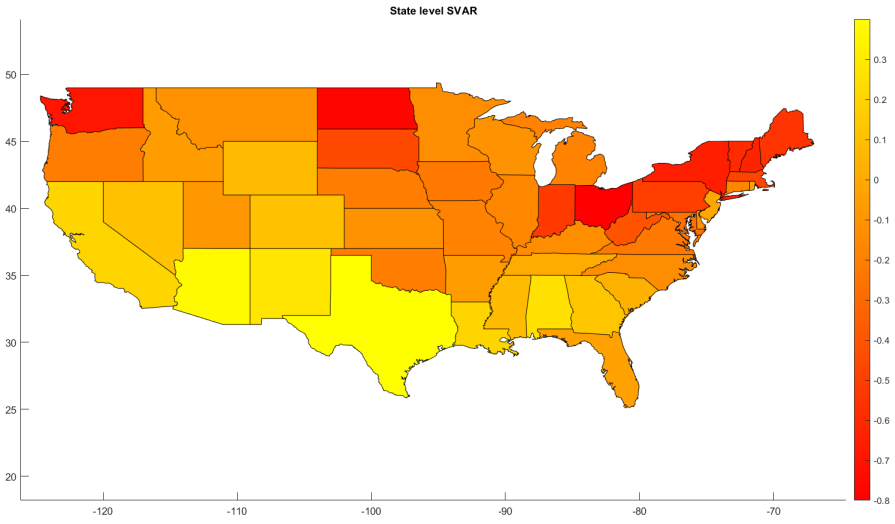


Figure (A1) presents the empirical measurement in September 2005 for the two temperature components. Figure (A) displays the relative state distribution of TD, figure (B) present and $TDVAR$ values. The heatmap presents in yellow states that are more exposed to the temperature statistics (higher anomalies or variability) and in red states less exposed.

However, aggregation at the U.S. level poses a second problem. regarding the weights associated at each state temperature deviation and deviation in variability. Following what we did with aggregation at the state level, the natural solution would be to employ an equally weighting method, $w_i = \frac{1}{N_s}$.

In such a way, however, an high $TD\text{-}VAR$ in a small state would have the same impact as the one experienced by a larger one, in the final index. The weighting method should thus reflect the importance each state exhibits in the nationwide index.

The two natural possibilities are GDP– and population-based weight. Both measures are not available at a monthly frequency, with the former defined quarterly and the latter yearly, so we forward feed the data series to obtain monthly frequency series comparable with $TD\text{-}VAR$ and TD, as described in Section ().

Employing the first approach, with GDP-weight, would imply relative importance of the most productive areas, meaning an economic specific impact of TD and $TD\text{-}VAR$. Applying the second approach instead would gather attention to what people perceive: more variability in a densely populated state is more important than the same event in a state less populated.

Figures A2 shows the difference in the U.S.-wide index for $TD\text{-}VAR$ according to the different weighting criteria. When comparing the three measures however there are minor differences just when considering equally weighting against population or GDP.

Appendix Figure A2. Different U.S. $TD\text{-}VAR$ acording to weighting method

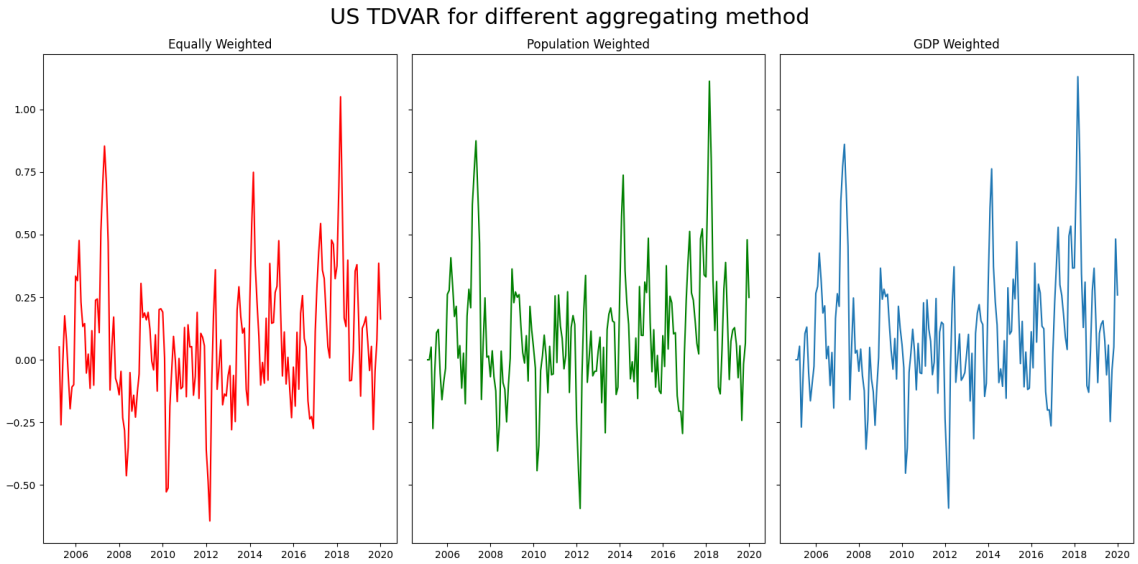


Figure (A2) shows three main index construction based on different weighting method of $us_t = \sum_i w_i * tdvar_{i,t}$ in the year 2005-2020. The left panel show equi-weighted index construction where $W_i = 1/N_i$ with $N_i = 50$. The central panel shows the population based weighting method where $w_i = pop_{i,t}/\sum_i pop_{i,t}$. In the right panel is plotted the US wide index weighted for state GDP, defining $w_i = \dfrac{GDP_{i,t}}{\sum_i GDP_{i,t}}$

Bibliography

Adams, Richard M, Cynthia Rosenzweig, Robert M Peart, Joe T Ritchie, Bruce A McCarl, J David Glyer, R Bruce Curry, James W Jones, Kenneth J Boote, and L Hartwell Allen, "Global climate change and US agriculture," *Nature*, 1990, *345* (6272), 219–224.

Addoum, Jawad M, David T Ng, and Ariel Ortiz-Bobea, "Temperature shocks and establishment sales," *The Review of Financial Studies*, 2020, *33* (3), 1331–1366.

_ , _ , and _ , "Temperature shocks and industry earnings news," *Available at SSRN 3480695*, 2021.

Alekseev, Georgij, Stefano Giglio, Quinn Maingi, Julia Selgrad, and Johannes Stroebel, "A quantity-based approach to constructing climate risk hedge portfolios," Technical Report, Working Paper 2021.

Alessandri, Piergiorgio and Haroon Mumtaz, "The macroeconomic cost of climate volatility," 2021.

Ardia, David, Keven Bluteau, Kris Boudt, and Koen Inghelbrecht, "Climate change concerns and the performance of green versus brown stocks," *National Bank of Belgium, Working Paper Research*, 2020, (395).

Augustin, Patrick, Linxiao Francis Cong, Ricardo Lopez Aliouchkin, and Roméo Tédongap, "Downside Risk and the Cross-section of Corporate Bond Returns," in "Proceedings of Paris December 2020 Finance Meeting EUROFIDAI-ESSEC" 2020.

Bai, Jennie, Turan G Bali, and Quan Wen, "Common risk factors in the cross-section of corporate bond returns," *Journal of Financial Economics*, 2019, *131* (3), 619–642.

Baiardi, Donatella and Mario Menegatti, "Pigouvian tax, abatement policies and uncertainty on the environment," *Journal of Economics*, 2011, *103* (3), 221–251.

Bansal, Ravi, Marcelo Ochoa, and Dana Kiku, "Climate change and growth risks," Technical Report, National Bureau of Economic Research 2017.

Barber, Brad M and Terrance Odean, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *The Review of Financial Studies*, 2008, *21* (2), 785–818.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman, "Can investors profit from the prophets? Security analyst recommendations and stock returns," *The Journal of Finance*, 2001, *56* (2), 531–563.

Bates, John M and Clive WJ Granger, "The combination of forecasts," *Journal of the Operational Research Society*, 1969, *20* (4), 451–468.

Battiston, Stefano, Antoine Mandel, Irene Monasterolo, Franziska Schütze, and Gabriele Visentin, "A climate stress-test of the financial system," *Nature Climate Change*, 2017, *7* (4), 283–288.

Benson, Charlotte and Edward J Clay, *Understanding the economic and financial impacts of natural disasters* number 4, World Bank Publications, 2004.

Benth, Fred Espen and Jūratė Šaltytė Benth, "The volatility of temperature and pricing of weather derivatives," *Quantitative Finance*, 2007, *7* (5), 553–561.

_ **and Juratė Šaltytė-Benth**, "Stochastic modelling of temperature variations with a view towards weather derivatives," *Applied Mathematical Finance*, 2005, *12* (1), 53–85.

Bentzen, Jan and Tom Engsted, "Short-and long-run elasticities in energy demand: a cointegration approach," *Energy Economics*, 1993, *15* (1), 9–16.

Bianchi, Daniele, Matthias Büchner, and Andrea Tamoni, "Bond risk premiums with machine learning," *The Review of Financial Studies*, 2021, *34* (2), 1046–1089.

Bigerna, Simona, "Estimating temperature effects on the Italian electricity market," *Energy Policy*, 2018, *118*, 257–269.

Black, Fischer and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy*, 1973, *81* (3), 637–654.

Blyth, William, Richard Bradley, Derek Bunn, Charlie Clarke, Tom Wilson, and Ming Yang, "Investment Risks under Uncertain Climate Change Policy," *Energy Policy*, 11 2007, *35*, 5766–5773.

Boissonnade, Auguste C, Lawrence J Heitkemper, and David Whitehead, "Weather data: cleaning and enhancement," *Climate Risk and the Weather Market*, 2002, pp. 73–98.

Bollerslev, Tim, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 1986, *31* (3), 307–327.

Bolton, Patrick and Marcin T Kacperczyk, "Carbon premium around the world," *Available at SSRN 3594188*, 2020.

Breiman, Leo, "Random forests," *Machine Learning*, 2001, *45* (1), 5–32.

Brieman, Leo, Jerome H Friedman, Richard A Olshen, and Charles J Stone, "Classification and regression trees," *Wadsworth Inc*, 1984, *67*.

Burke, Marshall and Vincent Tanutama, "Climatic constraints on aggregate economic output," Technical Report, National Bureau of Economic Research 2019.

_ , Solomon M Hsiang, and Edward Miguel, "Global non-linear effect of temperature on economic production," *Nature*, 2015, *527* (7577), 235–239.

Cai, Wenju, Simon Borlace, Matthieu Lengaigne, Peter Van Rensch, Mat Collins, Gabriel Vecchi, Axel Timmermann, Agus Santoso, Michael J McPhaden, Lixin Wu et al., "Increasing frequency of extreme El Niño events due to greenhouse warming," *Nature climate change*, 2014, *4* (2), 111–116.

Cai, Yongyang, Kenneth L Judd, Timothy M Lenton, Thomas S Lontzek, and Daiju Narita, "Environmental tipping points significantly affect the cost-benefit assessment of climate policies," *Proceedings of the National Academy of Sciences*, 2015, *112* (15), 4606–4611.

Calvin, Katherine, Marshall Wise, Leon Clarke, Jae Edmonds, Page Kyle, Patrick Luckow, and Allison Thomson, "Implications of simultaneously mitigating and adapting to climate change: initial experiments using GCAM," *Climatic Change*, 2013, *117* (3), 545–560.

Campbell, John Y and Glen B Taksler, "Equity volatility and corporate bond yields," *The Journal of Finance*, 2003, *58* (6), 2321–2350.

Campbell, Sean D and Francis X Diebold, "Weather forecasting for weather derivatives," *Journal of the American Statistical Association*, 2005, *100* (469), 6–16.

Cao, N, Valentina Galvani, and S Gubellini, "Firm-specific stock and bond predictability: New evidence from Canada," *International Review of Economics & Finance*, 2017, *51*, 174–192.

Cappiello, Lorenzo, Robert F Engle, and Kevin Sheppard, "Asymmetric dynamics in the correlations of global equity and bond returns," *Journal of Financial Econometrics*, 2006, *4* (4), 537–572.

Carhart, Mark M, "On persistence in mutual fund performance," *The Journal of Finance*, 1997, *52* (1), 57–82.

Carr, Peter and Liuren Wu, "Decomposing long bond returns: A decentralized modeling approach," *Baruch College Zicklin School of Business Research Paper*, 2019, (2019-08), 06.

Castellano, Rosella, Roy Cerqueti, and Giulia Rotundo, "Exploring the financial risk of a temperature index: A fractional integrated approach," *Annals of Operations Research*, 2020, *284* (1), 225–242.

Cavanagh, Ralph, Ashok Gupta, Dan Lashof, and Marika Tatsutani, "Utilities and CO2 emissions: Who bears the risks of future regulation?," *The Electricity Journal*, 1993, *6* (2), 64–75.

Ceglar, Andrej, Andrea Toreti, Rémi Lecerf, Marijn Van der Velde, and Frank Dentener, "Impact of meteorological drivers on regional inter-annual crop yield variability in France," *Agricultural and forest meteorology*, 2016, *216*, 58–67.

Chang, Yoosoon, Chang Sik Kim, J Isaac Miller, Joon Y Park, and Sungkeun Park, "A new approach to modeling the effects of temperature fluctuations on monthly electricity demand," *Energy Economics*, 2016, *60*, 206–216.

Chen, Linda H and Lucia Silva Gao, "The pricing of climate risk," *Journal of Financial and Economic Practice, Vol12 (2), Spring*, 2012, pp. 115–131.

Choi, Darwin, Zhenyu Gao, and Wenxi Jiang, "Attention to global warming," *The Review of Financial Studies*, 2020, *33* (3), 1112–1145.

Clarkson, Peter M, Yue Li, Gordon D Richardson, and Florin P Vasvari, "Does it really pay to be green? Determinants and consequences of proactive environmental strategies," *Journal of Accounting and Public Policy*, 2011, *30* (2), 122–144.

Cochrane, John H. and Monika Piazzesi, "Bond Risk Premia," *American Economic Review*, March 2005, *95* (1), 138–160.

Colacito, Riccardo, Bridget Hoffmann, and Toan Phan, "Temperature and growth: A panel analysis of the United States," *Journal of Money, Credit and Banking*, 2019, *51* (2-3), 313–368.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, "In search of attention," *The Journal of Finance*, 2011, *66* (5), 1461–1499.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, 2012, *4* (3), 66–95.

Deschênes, Olivier and Michael Greenstone, "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather," *American Economic Review*, 2007, *97* (1), 354–385.

Diaz, Delavane and Frances Moore, "Quantifying the economic risks of climate change," *Nature Climate Change*, 2017, *7* (11), 774–782.

Dick-Nielsen, Jens, Peter Feldhütter, and David Lando, "Corporate bond liquidity before and after the onset of the subprime crisis," *Journal of Financial Economics*, 2012, *103* (3), 471–492.

Diebold, Francis X and Glenn D Rudebusch, "On the evolution of US temperature dynamics," in "Essays in Honor of M. Hashem Pesaran: Prediction and Macro Modeling," Emerald Publishing Limited, 2022.

Dietz, Simon and Nicholas Stern, "Endogenous growth, convexity of damage and climate risk: how Nordhaus' framework supports deep cuts in carbon emissions," *The Economic Journal*, 2015, *125* (583), 574–620.

Dixon, John K, "Pattern recognition with partly missing data," *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, *9* (10), 617–621.

Donadelli, Michael, Marcus Jüppner, Antonio Paradiso, and Christian Schlag, "Temperature volatility risk," *University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No*, 2019, *5*.

_ , _ , _ , and _ , "Computing macro-effects and welfare costs of temperature volatility: a structural approach," *Computational Economics*, 2020, pp. 1–48.

_ , _ , Max Riedel, and Christian Schlag, "Temperature shocks and welfare costs," *Journal of Economic Dynamics and Control*, 2017, *82*, 331–355.

Dornier, Fabien and Michel Queruel, "Caution to the wind," *Energy & Power Risk Management*, 2000, *13* (8), 30–32.

Downie, John and Wendy Stubbs, "Evaluation of Australian companies' scope 3 greenhouse gas emissions assessments," *Journal of Cleaner Production*, 2013, *56*, 156–163.

Duan, Tinghua, Frank Weikai Li, and Quan Wen, "Is Carbon Risk Priced in the Cross Section of Corporate Bond Returns?," *Available at SSRN 3709572*, 2020.

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, "Hedging climate change news," *The Review of Financial Studies*, 2020, *33* (3), 1184–1216.

Fabozzi, Frank J and Steven V Mann, *The handbook of fixed income securities*, McGraw-Hill Education, 2012.

Faccini, Renato, Rastin Matin, and George S Skiadopoulos, "Dissecting Climate Risks: Are they Reflected in Stock Prices?," *Available at SSRN 3795964*, 2021.

Fama, Eugene F, "Stock returns, expected returns, and real activity," *The Journal of Finance*, 1990, *45* (4), 1089–1108.

_ and R Kenneth, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 1993, *33* (1), 3–56.

Fankhauser, Samuel and Thomas KJ McDermott, "Understanding the adaptation deficit: why are poor countries more vulnerable to climate events than rich countries?," *Global Environmental Change*, 2014, *27*, 9–18.

French, Kenneth R., "Fama French Repository," 2020.

French, Kenneth R, G William Schwert, and Robert F Stambaugh, "Expected stock returns and volatility," *Journal of Financial Economics*, 1987, *19* (1), 3–29.

Friedman, Jerome H, *The elements of statistical learning: Data mining, inference, and prediction*, Springer Open, 2017.

Gilkeson, James H and Stephen D Smith, "The convexity trap: pitfalls in financing mortgage portfolios and related securities," *Economic Review-Federal Reserve Bank of Atlanta*, 1992, *77* (6), 14.

Gleason, Terry C and Richard Staelin, "A proposal for handling missing data," *Psychometrika*, 1975, *40* (2), 229–252.

Green, Jessica F, "Does carbon pricing reduce emissions? A review of ex-post analyses," *Environmental Research Letters*, 2021.

Greenwood, Robin and Samuel G Hanson, "Issuer quality and corporate bond returns," *The Review of Financial Studies*, 2013, *26* (6), 1483–1525.

Griffin, Paul A, David H Lont, and Estelle Y Sun, "The relevance to investors of greenhouse gas emission disclosures," *Contemporary Accounting Research*, 2017, *34* (2), 1265–1297.

Grippa, Pierpaolo, Jochen Schmittmann, and Felix Suntheim, "Climate Change and Financial Risk: Central banks and financial regulators are starting to factor in climate change," *Finance & Development*, 2019, *56* (004).

Gros, Daniel, Philip R Lane, Sam Langfield, Sini Matikainen, Marco Pagano, Dirk Schoenmaker, and Javier Suarez, *Too late, too sudden: Transition to a low-carbon economy and systemic risk* number 6, Reports of the Advisory Scientific Committee, 2016.

Hansen, James, Donald Johnson, Andrew Lacis, Sergej Lebedeff, Pius Lee, David Rind, and Gary Russell, "Climate impact of increasing atmospheric carbon dioxide," *Science*, 1981, *213* (4511), 957–966.

_ , **Makiko Sato, and Reto Ruedy**, "Perception of climate change," *Proceedings of the National Academy of Sciences*, 2012, *109* (37), E2415–E2423.

Hassler, John and Per Krusell, "Economics and climate change: integrated assessment in a multi-region world," *Journal of the European Economic Association*, 2012, *10* (5), 974–1000.

He, Xin, Guanhao Feng, Junbo Wang, and Chunchi Wu, "Predicting individual corporate bond returns," *Available at SSRN*, 2021.

Hertwich, Edgar G and Richard Wood, "The growing importance of scope 3 greenhouse gas emissions from industry," *Environmental Research Letters*, 2018, *13* (10), 104013.

Hjort, Ingrid, "Potential climate risks in financial markets: A literature overview," Technical Report, Memorandum 2016.

Höck, André, Christian Klein, Alexander Landau, and Bernhard Zwergel, "The effect of environmental sustainability on credit risk," *Journal of Asset Management*, 2020, pp. 1–9.

Hong, Harrison, G Andrew Karolyi, and José A Scheinkman, "Climate finance," *The Review of Financial Studies*, 2020, *33* (3), 1011–1023.

Hong, Yongmiao, Hai Lin, and Chunchi Wu, "Are corporate bond market returns predictable?," *Journal of Banking & Finance*, 2012, *36* (8), 2216–2232.

Hope, Chris, "The PAGE09 integrated assessment model: A technical description," *Cambridge Judge Business School Working Paper*, 2011, *4* (11).

Hsiang, Solomon, "Climate econometrics," *Annual Review of Resource Economics*, 2016, *8*, 43–75.

_ **and Robert E Kopp**, "An economist's guide to climate change science," *Journal of Economic Perspectives*, 2018, *32* (4), 3–32.

Huynh, Thanh and Ying Xia, "Climate change news risk and corporate bond returns," *Journal of Financial and Quantitative Analysis, forthcoming*, 2021.

Ilhan, Emirhan, Zacharias Sautner, and Grigory Vilkov, "Carbon tail risk," *The Review of Financial Studies*, 2021, *34* (3), 1540–1571.

Ilmanen, Antti, "Time-varying expected returns in international bond markets," *The Journal of Finance*, 1995, *50* (2), 481–506.

_ , "Stock-bond correlations," *The Journal of Fixed Income*, 2003, *13* (2), 55–66.

Israel, Ronen, Diogo Palhares, and Scott A Richardson, "Common factors in corporate bond returns," *Journal of Investment Management*, 2017.

Jewson, Stephen and Anders Brix, *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*, Cambridge University Press, 2005.

Karl, Thomas R, Henry F Diaz, and George Kukla, "Urbanization: Its detection and effect in the United States climate record," *Journal of climate*, 1988, *1* (11), 1099–1123.

Kolk, Ans and Jonatan Pinkse, "Business Responses to Climate Change: Identifying Emergent Strategies," *California Management Review*, 04 2005, *47*.

Kotz, Maximilian, Leonie Wenz, Annika Stechemesser, Matthias Kalkuhl, and Anders Levermann, "Day-to-day temperature variability reduces economic growth," *Nature Climate Change*, 2021, *11* (4), 319–325.

Kumar, Alok, Wei Xin, and Chendi Zhang, "Climate sensitivity and predictable returns," *Available at SSRN 3331872*, 2019.

Lebelle, Martin, Souad Lajili Jarjir, and Syrine Sassi, "Corporate green bond issuances: An international evidence," *Journal of Risk and Financial Management*, 2020, *13* (2), 25.

Lee, Su-Yol, Yun-Seon Park, and Robert D Klassen, "Market responses to firms' voluntary climate change information disclosure and carbon communication," *Corporate Social Responsibility and Environmental Management*, 2015, *22* (1), 1–12.

Lewandowski, Stefan, "Corporate carbon and financial performance: The role of emission reductions," *Business Strategy and the Environment*, 2017, *26* (8), 1196–1211.

Lewis, Sophie C and Andrew D King, "Evolution of mean, variance and extremes in 21st century temperatures," *Weather and Climate Extremes*, 2017, *15*, 1–10.

Liaw, Andy, Matthew Wiener et al., "Classification and regression by randomForest," *R news*, 2002, *2* (3), 18–22.

Lin, Hai, Chunchi Wu, and Guofu Zhou, "Forecasting corporate bond returns with a large set of predictors: An iterated combination approach," *Management Science*, 2018, *64* (9), 4218–4238.

_ , Junbo Wang, and Chunchi Wu, "Liquidity risk and expected corporate bond returns," *Journal of Financial Economics*, 2011, *99* (3), 628–650.

Makridis, C, "Can you feel the heat? Extreme temperatures, stock returns, and economic sentiment," *Social Science Research Network: Rochester, NY, USA*, 2018.

Meucci, Attilio, *Risk and asset allocation*, Vol. 1, Springer, 2005.

Moore, Frances C and Delavane B Diaz, "Temperature impacts on economic growth warrant stringent mitigation policy," *Nature Climate Change*, 2015, *5* (2), 127–131.

Nembrini, Stefano, Inke R König, and Marvin N Wright, "The revival of the Gini importance?," *Bioinformatics*, 2018, *34* (21), 3711–3718.

Neumann, James E and Kenneth Strzepek, "State of the literature on the economic impacts of climate change in the United States," *Journal of Benefit-Cost Analysis*, 2014, *5* (3), 411–443.

Nordhaus, William, "Estimates of the social cost of carbon: concepts and results from the DICE-2013R model and alternative approaches," *Journal of the Association of Environmental and Resource Economists*, 2014, *1* (1/2), 273–312.

_ **and Paul Sztorc**, "DICE 2013R: Introduction and User's Manual," 2013.

Nordhaus, William D, "Economic growth and climate: the carbon dioxide problem," *The American Economic Review*, 1977, pp. 341–346.

Nozawa, Yoshio, *Corporate bond premia*, The University of Chicago, 2013.

Pankratz, Nora and Christoph Schiller, "Climate change and adaptation in global supply-chain networks," in "Proceedings of Paris December 2019 Finance Meeting EUROFIDAI-ESSEC, European Corporate Governance Institute–Finance Working Paper" number 775 2021.

Pastor, Lubos, Robert F Stambaugh, and Lucian A Taylor, "Dissecting Green Returns," Technical Report, National Bureau of Economic Research 2021.

Pogue, Thomas F. and Robert M. Soldofsky, "What's in a Bond Rating," *Journal of Financial and Quantitative Analysis*, 1969, *4* (2), 201–228.

Qian, Wei, Ani Suryani, and Ke Xing, "Does carbon performance matter to market returns during climate policy changes? Evidence from Australia," *Journal of Cleaner Production*, 2020, p. 121040.

Quayle, Robert G and Henry F Diaz, "Heating degree day data applied to residential heating energy consumption," *Journal of Applied Meteorology and Climatology*, 1980, *19* (3), 241–246.

Rapach, David E, Jack K Strauss, and Guofu Zhou, "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy," *The Review of Financial Studies*, 2010, *23* (2), 821–862.

Reboredo, Juan C and Andrea Ugolini, "Price connectedness between green bond and financial markets," *Economic Modelling*, 2020, *88*, 25–38.

Ren, Guoyu, Yaqing Zhou, Ziying Chu, Jiangxing Zhou, Aiying Zhang, Jun Guo, and Xuefeng Liu, "Urbanization effects on observed surface air temperature trends in North China," *Journal of Climate*, 2008, *21* (6), 1333–1348.

Rohde, Robert A and Zeke Hausfather, "The Berkeley Earth land/ocean temperature record," *Earth System Science Data*, 2020, *12* (4), 3469–3479.

Rohde, Robert, Richard Muller, Robert Jacobsen, Saul Perlmutter, Arthur Rosenfeld, Jonathan Wurtele, Judith Curry, Charlotte Wickham, and Steven Mosher, "Berkeley earth temperature averaging process," *Geoinformatics & Geostatistics: An Overview*, 2013, *1* (2), 1–13.

Rubin, Donald B, "Inference and missing data," *Biometrika*, 1976, *63* (3), 581–592.

Russo, Michael V and Amy Minto, "Competitive strategy and the environment: A field of inquiry emerges," in "The Oxford handbook of business and the natural environment" 2012.

Sautner, Zacharias, Laurence van Lent, Grigory Vilkov, and Ruishen Zhang, "Firm-level climate change exposure," *European Corporate Governance Institute–Finance Working Paper*, 2020, (686).

Scheinkman, Jose A and Blake LeBaron, "Nonlinear dynamics and stock returns," *Journal of business*, 1989, *62* (3), 311–337.

Schlenker, Wolfram and Charles A Taylor, "Market expectations of a warming climate," *Journal of Financial Economics*, 2021, *142* (2), 627–640.

Schneider, Thomas E, "Is environmental performance a determinant of bond pricing? Evidence from the US pulp and paper and chemical industries," *Contemporary Accounting Research*, 2011, *28* (5), 1537–1561.

Son, Hyojoo and Changwan Kim, "Short-term forecasting of electricity demand for the residential sector using weather and social variables," *Resources, conservation and recycling*, 2017, *123*, 200–207.

Starr, Martha, "The effects of weather on retail sales," *Available at SSRN 221728*, 2000.

Stern, Nicholas, "The economics of climate change," *American Economic Review*, 2008, *98* (2), 1–37.

Stock, James H and Mark W Watson, "A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series," 1998.

Titman, Sheridan, KC John Wei, and Feixue Xie, "Capital investments and stock returns," *Journal of financial and Quantitative Analysis*, 2004, *39* (4), 677–700.

Tol, Richard and David and Anthoff, "Fund Model Repository," 2014.

Tol, Richard SJ, "The marginal costs of greenhouse gas emissions," *The Energy Journal*, 1999, *20* (1), 61–81.

Toohey, Matthew, Kirstin Krüger, Hauke Schmidt, Claudia Timmreck, Michael Sigl, Markus Stoffel, and Rob Wilson, "Disproportionately strong climate forcing from extratropical explosive volcanic eruptions," *Nature Geoscience*, 2019, *12* (2), 100–107.

Tsai, Hui-Ju, "The informational efficiency of bonds and stocks: The role of institutional sized bond trades," *International Review of Economics & Finance*, 2014, *31*, 34–45.

Xu, Yangyang, Veerabhadran Ramanathan, and David G Victor, "Global warming will happen faster than we think," *Nature*, 2018, *564* (7734), 30–32.

Yang, Ming, William Blyth, Richard Bradley, Derek Bunn, Charlie Clarke, and Tom Wilson, "Evaluating the power investment options with uncertainty in climate policy," *Energy Economics*, 2008, *30* (4), 1933–1950.

Zachariadis, Theodoros and Nicoletta Pashourtidou, "An empirical analysis of electricity consumption in Cyprus," *Energy Economics*, 2007, *29* (2), 183–198.

Zanobetti, Antonella, Marie S O'Neill, Carina J Gronlund, and Joel Schwartz, "Temperature Variability and long-term survival," in "ISEE Conference Abstracts" 2011.

Zerbib, Olivier David, "The green bond premium," *Available at SSRN 2890316*, 2017.

_ , "The effect of pro-environmental preferences on bond prices: Evidence from green bonds," *Journal of Banking & Finance*, 2019, *98*, 39–60.

Zhang, Lei and Min Zhu, "Corporate Exposure to Weather and Bond Yield Spread," *Available at SSRN 3547895*, 2021.

Zivin, Joshua Graff and Matthew Neidell, "Temperature and the allocation of time: Implications for climate change," *Journal of Labor Economics*, 2014, *32* (1), 1–26.