LUISS

Department of Economics and Finance

Chair of Econometric Theory

THE ROLE OF SECONDARY MARKETS ON CARBON EMISSIONS: AN EMPIRICAL ANALYSIS

Supervisor:
Prof. Paolo Santucci De Magistris

Co-Supervisor:
Prof. Federico Carlini

Candidate:
Giampietro Molaioni
ID: 738461

Academic year 2021/2022

Abstract

The research discusses the relationship between industrial emissions of greenhouse gases (GHG) and financial markets, and how the latter can play a key role in reducing or increasing global carbon emissions. the study focuses on the effects that companies' GHG emissions can have on future cashflows and how this can impact stock market prices and returns. If this effect is large enough and statistically significant, GHG emissions are able to influence companies cost of capital and thus management decisions. Unlike other studies, this research makes no assumptions about the presence of different types of investors in the market, is fully empirical and uses Two Stage Least Squares regression to avoid estimation bias due to reverse causality between returns and GHG emissions

Contents

Abbreviations and Nomenclature

CAPM – Capital Asset Pricing Model

CESM – Community Earth System Model

CO2 – Carbon Dioxide

CO2e – Carbon Dioxide equivalent.

CH4 – Methane

CMIP – Coupled Model Intercomparison Project

DEMETER – Decarbonisation Model for Endogenous Technologies Emission Reductions

DICE – Dynamic Integrated Climate-Economy model

EDGAR – Emissions Database for Global Atmospheric Research

ESG – Environmental, Social, and Governance

GDP – Gross Domestic Product

GHG – Greenhouse Gases

GMM – Generalized Method of Moments

HadGEM – Hadley Centre Global Environmental Model

kWh – Kilowatt-hour

IAM – Integrated assessment models

IV – Instrumental Variable

IPCC – Intergovernmental Panel on Climate Change

M&O – Maintainance and Operations

N2O – Nitrous Oxide

NASA – National Areonautic Space Administration

NCAR – National Center for Atmospheric Reserch

OECD – Organization for Economic Co-operation and Development

OLS – Ordinary Least Squares

R&D – Research and Development

R&DICE – Research and Development Integrated Climate-Economy model

REMIND – Regional Model of Investment and Development

TSCS – Two Stage Cross-Sectional

TSLS – Two Stages Least Squares

USD – United States dollars

WCRP – World Climate Research Programme

Chapter 1

Introduction

Between 2012 and 2021 S&P 500[1] and STOXX 600[2] market capitalization increased by 10,62% and 10,92% yearly, while the absolute emissions of companies that were publishing Carbon Dioxide equivalent [3] (CO2e) emissions in the two indexes from 2012 decreased by 4,93% and 4,80% yearly, respectively (Refintiv, 2021). Over the same period, absolute CO2e emissions in the United States and Europe decreased by 0,02% and 2,35%, leading to 6 and 3 billion tonnes of CO2e emitted in each territory (OECD, 2021). the research explained how industrial air pollution is the first driver of climate change which is the cause of the rising global temperature and sea level (IPCC, 2007). Nordhaus (2018) explains how Greenhouse gases (GHG) remain partially trapped in the atmosphere for short-term, long-term, and permanent. GHG in the atmosphere keep solar radiation longer over the globe, preventing them from going out into space. Higher radiation results in higher temperatures, which also contribute to polar ice melting, and thus rising sea levels. Rising sea and temperature levels generate a series of climate disasters such as the disappearance of some islands, severe droughts, famines, wildfires, and extreme weather events generated by storms such as floods and structure damages. the board of climate experts created to inform policymakers globally, named the Intergovernmental Panel on Climate Change (IPCC), explained how, if we want to mitigate and contain those disasters, we should reduce CO2e global emissions to keep the average global warming under 2°, preferably 1.5° if we want to avoid entire island states being covered by the water (IPCC, 2021). We will refer to *carbon emissions* as the atmospheric emissions of any GHG expressed in CO2e units for the purposes of this study. Since industrial

[1] S&P 500 is a stock index made with 500 leading publicly traded companies in the United States, and it is the representative stock index of the United States market

[2] STOXX 600 is a stock index made with 600 leading publicly traded companies in Europe, and it is the representative stock index of the European market

[3] It is a measurement of the potential for greenhouse gases to cause global warming. It is used to compare the global warming effects of the emissions of various GHG: for example 1 tonne of CO2 = 1 tonne CO2e, while 1 tonne of Nitrous Oxide = 320 tonnes CO2e

emissions of GHG and their effect on solar radiation is the main driver of global warming, financial markets can play a key role in increasing or reducing global *carbon emissions*.

Financial markets can embrace industry and company changes by discounting the effects that some companies' features can have on future cashflows, reducing stock market prices today. As managers and shareholders are generally stock price maximizers, they tend to adapt the projects of the company to mitigate negative externalities and maximize profits (Harris, 2022). Relatively to *carbon emissions*, the negative impact on future cashflows can be generated by regulations restricting certain practices, carbon tax, and competition with technological innovators. All these future externalities represent a risk for a company emitting GHG. Riskier future cash flows are discounted at a higher cost of capital, represented by the weighted sum of the cost of equity and debt (Cheema-Fox *et al.*, 2019). A higher cost of capital for the company means that discounted future cashflows are lower, and so it is the fundamental stock price. This higher cost of capital should push managers to pursue projects that mitigate *carbon emissions* to reduce the cost of capital when the cost of mitigating *carbon emissions* is lower than the increased costs of capital due to the higher *carbon emissions* (Schmidt, 2014).

Investing in projects that have lower *carbon intensity* than projects in the same industry, can instead reduce the cost of capital of a company and specifically the cost with which that project is financed. *Carbon intensity* is defined as the tonnes of Carbon dioxide (CO2) emissions over 1 million United States dollars (USD) in revenues generated by an investment project or a company (Refinitiv, 2021). Financing a project with higher *carbon intensity* than the projects with which it competes, increases the *carbon intensity* of the industry, and may even reduce its absolute level of *carbon emissions*. Investing in projects that are less carbon intensive than the ones in the same industry, instead, reduces the *carbon intensity* of that industry and may also reduce its absolute *carbon emissions* (H. Hong and M. Kacperczyk, 2009).

Since managers and shareholders are generally stock price maximizers, they take decisions to maximize companies' profits. They do this either by increasing the scale of the activity or by launching new businesses. When these activities directly or indirectly emit GHG, these emissions grow together with profits.

This research tries to answer the following questions: Relatively to secondary markets, how does a company's *carbon emissions* level influence its cost of capital? How can we observe the effect on the cost of capital by isolating the increase in *carbon emissions* from the increase in revenues?

To study the phenomena that regard investors' climate change awareness, researchers rely on very different types of studies, but all of them can be included in the two following streams: Research is based either on economic equilibrium models to assess the presence of impact investors (Harris 2022, T. De Angelis *et al.* 2021), or on asset pricing models that assess ESG products features or ESG preferences and rely mostly on ESG scores or variables built through machine learning tools (J.B. Berk, and J.H. van Binsbergen, 2021). Sometimes the results of these models are contradictive. the model presented in the following research relies only on empirical data and on the Two Stages Least Squares (TSLS) estimation with an Instrumental Variable to overcome endogeneity biases.

The presented model is purely empirical and relies on a TSLS estimation to avoid errors of a model that otherwise would suffer from endogeneity bias (James and Singh, 1978). the model may suffer from reverse causality, which is a mechanism that the research tries to explain: *carbon emissions* and the related externalities may influence the cost of capital, and this in turn may influence the level of *carbon emissions*.

Chapter 2

Industry Contribution to Climate

Human activity has generated significant global change, particularly through the emission of GHG into the atmosphere. the largest contributor to GHG emissions is the industrial sector, primarily by burning fossil fuels for energy production.

According to the IPCC, in 2018, the industrial sector accounted for approximately 31% of total GHG emissions globally. Within the industrial sector, the consumption and production of energy, specifically through the burning of natural gas, coal, and oil accounted for nearly 75% of these emissions (IPCC, 2021).

In addition to energy production, the industrial sector also emits GHG through processes such as deforestation and land-use change, chemical production, and waste management.

The impact of industrial *carbon emissions* on the global climate can be seen through the increasing concentrations of atmospheric CO2 and other GHG, which trap heat and warm the planet. This leads to a range of environmental and socio-economic impacts, including rising global temperatures, melting glaciers, sea level rise, more intense and frequent natural disasters, and changes to ecosystems and biodiversity (World Bank, 2021).

To mitigate the impacts of human-generated global change, it is crucial to reduce industrial GHG emissions through transitioning to cleaner, low-carbon energy sources, increasing energy efficiency, and reducing waste and deforestation (IPCC, 2007).

There is a large body of scientific evidence that confirms the role of industrial emissions in contributing to climate change. This evidence comes from a variety of sources, including direct measurements of GHG emissions and atmospheric concentrations, climate models, and observational data on the impacts of climate change.

One key line of evidence is the increase in atmospheric concentrations of GHG, particularly CO2, which is largely caused by human activities, particularly the burning of fossil fuels for energy production. the IPCC reports that CO2 concentrations have increased by about 40% since pre-industrial times and are continuing to rise (IPCC, 2018).

Another important piece of evidence is the correlation between the increase in GHG emissions and the observed warming of the planet (IPCC, 2014).

Climate models

Climate models have consistently shown that the warming trend observed in recent decades can only be explained by taking into account the increase in GHG emissions caused by human activities.

Additionally, observational data has demonstrated the effects of climate change, including rising sea levels, melting glaciers, and changes to ecosystems and biodiversity (Shaftel *et al.*, 2021). These impacts are consistent with the predictions of climate models and the expected consequences of global warming.

CMIP

The Coupled Model Intercomparison Project (CMIP) is a collaborative effort between the international climate modeling community and the World Climate Research Programme (WCRP) to evaluate and compare the performance of climate models. In CMIP, simulations are run using multiple climate models, allowing for a systematic comparison of model performance and the assessment of uncertainties in future climate projections.

CMIP is a protocol for coordinating and comparing the results of climate model simulations conducted by multiple climate modeling groups around the world. the goal of CMIP is to provide a standardized set of simulations that can be used to assess and compare the performance of different climate models, and to evaluate the models' ability to simulate the Climate system of the Earth and its response to changes in atmospheric composition (Climate Model Intercomparison Project, 2021).

In CMIP simulations, climate models are run under a range of scenarios, including historical conditions and future projections of GHG emissions. the results of these simulations are then compared and analyzed to evaluate the models' performance and to assess the response of the Climate system of the Earth to different GHG emission scenarios.

The results of CMIP simulations have consistently shown that the observed warming trend in recent decades is extremely unlikely to be due to natural climate variability alone and can be explained only by the increase in GHG emissions caused by human activities (IPCC, 2013).

The CMIP models have also provided robust projections of future climate change, including an increase in global mean surface temperature, rising sea levels, and changes in precipitation patterns. These projections have been used to inform policy and decision-making at the national and international level (IPCC, 2014).

CESM

A comprehensive and integrated computer model of the Climate system of the Earth is called the Community Earth System Model (CESM). Scientists from all around the world use it to study and comprehend the Earth's climate and how it responds to changes in the seas and atmosphere. It was created and is maintained by the National Center for Atmospheric Research (NCAR).

The CESM depicts all of the interactions between the ocean, atmosphere, sea ice, and land surface and the exchange of gases and energy between the atmosphere and the ocean. the model includes information on past and present climatic circumstances, including temperature, precipitation, and air composition observations. It is based on the physical, chemical, and biological laws that govern how the Climate system of the Earth behaves (NCAR, 2021).

The model is run in CESM simulations under different conditions, such as historical conditions and projected GHG emission levels in the future. Based on these simulations, scientists may predict future impacts of climate change and assess how the Climate system of the Earth will behave under various GHG emission scenarios (IPCC, 2014).

The outcomes of CESM models have repeatedly demonstrated that the observed warming trend started from the end of the last century cannot possibly be explained by natural climate variability alone but rather must be attributed to the rise in GHG emissions brought on by human activity. the results are in accordance with numerous lines of evidence, like the observed rise in atmospheric GHG concentrations, the seen warming of the oceans, and the observed melting of the polar ice caps (IPCC, 2014).

CESM is a particular climate model created by NCAR, whereas CMIP is a protocol for coordinating and comparing the outcomes of simulations carried out by other climate modeling groups. While CMIP is intended to assess and compare the performance of various climate models, CESM is employed to simulate the Climate system of the Earth and its reaction to changes in atmospheric composition (NCAR, 2021).

HadGEM

The Hadley Centre for Climate Change at the UK's Met Office created the Hadley Centre Global Environmental Model (HadGEM), a climate model. HadGEM is used, like other climate models, to mimic the Climate system of the Earth and forecast future climatic changes.

HadGEM begins the modeling process by simulating the processes that take place on Earth, including those that affect the ocean, atmosphere, sea ice, and land surface. the model integrates information on both current and historical climate conditions, and the representations of the passage of energy, water, and gases between the atmosphere and the ocean (Collins *et al.* 2010).

HadGEM is engineered in way that makes it more computationally efficient than other climate models, like CESM and CMIP, which is one of their main differences. This enables to operate the model on different computer systems, from powerful supercomputers to basic desktop computers (Jones *et al*., 2012).

HadGEM is made to have a higher geographic resolution, which is another distinction it has over other climate models. As a result, the model can more accurately capture details of smaller-scale features and processes, as individual storms (Jones *et al.*, 2012).

HadGEM is used in a variety of scenarios, including past circumstances and future GHG emission forecasts, just like other climate models. Predictions on how the Climate system of the Earths will react to various GHG emission scenarios are then based on the outcomes of these simulations.

MIROC

A complex climate model, the Model for Interdisciplinary Research on Climate (MIROC), mimics how the Climate system of the Earth reacts to various environmental conditions, such as GHG emissions. the model depicts the ocean, atmosphere, sea ice, and land surface as well as the flow of gases, water, and energy between these elements (Watanabe *et al*., 2011).

Beginning with the entry of data on GHG emissions, land use, and other environmental variables, such as information on past and present climate conditions, the modeling process in MIROC can begin. Following that, the model emulates the natural processes, like atmospheric circulation, ocean currents, and the movement of energy, water, and gases between the atmosphere and the ocean, that control the behaviour of the climate on the Earth.

MIROC employs a non-hydrostatic atmospheric model, which is one of its main distinctions from other climate models like CESM, CMIP, and HadGEM. By not assuming that the atmosphere is in hydrostatic equilibrium, the model enables more accurate simulation of small-scale atmospheric processes as deep convection and storms (Watanabe *et al*., 2012).

The non-hydrostatic equilibrium refers to the assumption that the atmosphere is not in a state of hydrostatic equilibrium, meaning that the atmosphere's pressure and density are not in balance with the gravitational force acting on the atmosphere. In traditional climate models, the assumption of hydrostatic equilibrium allows to simplify the mathematical equations that describe atmosphere's behaviour. However, this assumption can lead to errors in the simulation of certain atmospheric processes, such as deep convection and storms, that are important for the Earth's climate (Watanabe *et al.*, 2011).

The term "non-hydrostatic equilibrium" describes the idea that the atmosphere isn't in a condition of hydrostatic equilibrium, meaning that its pressure and density aren't in a balance that's proportional with the gravitational force operating on it. the assumption of hydrostatic equilibrium is employed in conventional climate models to streamline the mathematical equations describing the atmosphere's behaviour. However, this supposition may result in mistakes when simulating key atmospheric processes that are crucial to the Earth's climate, like deep convection and storms (Watanabe *et al.*, 2011).

The model, for instance, can simulate changes in vegetation cover and land use brought on by human activities like agriculture and deforestation, and it can forecast the effects these changes will have on the Earth's climate, such as adjustments to temperature, precipitation, and atmospheric carbon dioxide levels.

The outcomes of the MIROC simulations are then utilised, as in the other instances, to assess the model's effectiveness and create forecasts regarding the future climate of the Earth, including temperature, precipitation, and other important factors. These projections are used to guide policy and decision-making to mitigate and adapt to climate change and its effects, because they are computed on several scenarios of GHG emissions and other environmental conditions (IPCC, 2018).

Integrated Assessment Models

While climate models study the impacts of possible scenarios of *carbon emissions* on the earth, there are other models that forecast *carbon emissions* depending on socioeconomic scenarios, and their subsequent effect on earth's average temperature. To model the *carbon emissions* generated by human activity, the Integrated Assessment Models (IAM) are used by academics. IAM knowledge from more than one domain into a single framework to answer to research questions or model phenomena that require multidisciplinary comprehension. These models are thus suited for a complex subject such as climate change (Nordhaus, 2019).

DICE

Nordhaus (2016) presents the latest published Dynamic Integrated Climate-Economy (DICE) IAM. In this framework Uncontrolled industrial CO2 emissions are given by a level of *carbon intensity* or the CO2- output ratio, $\sigma(t)$, times gross output. Total CO2 emissions, $E(t)$, are equal to uncontrolled emissions reduced by the emissions-reduction rate, μ(t), plus exogenous land-use emissions.

$$E(t) = \sigma(t)[1 - \mu(t)]Y(t) + E^l(t) \qquad \text{(1-DICE)}$$

According to the DICE model, global *carbon intensity* decays at the empirical historical rate, which is 1,5% each year.

$$\sigma(t) = \sigma(0)exp(-\upsilon t), \quad \upsilon = 1{,}5\%. \qquad \text{(2-DICE)}$$

The emission reduction rate depends on the substitution between labour, capital and carbon energy, which is energy produced with fossil fuel combustion. This model, thus, depends on the hypothesis that a rise in the price of carbon energy will force economic agents to use more capital and labour rather than carbon energy to obtain this economic global output. This model takes the future path of *carbon intensity* as a simple function of time while the consumption of *carbon emissions* is an endogenous variable depending on *carbon emissions*' price (Nordhaus, 2016).

R&DICE

The Research and Development Integrated Climate-Economy model (R&DICE) model (Nordhaus, 2014) contributes to the climate change IAM literature, adding an endogenous model to forecast *carbon emissions* dependending on technology investments. Technology change has an impact on *carbon intensity* in the R&DICE model. In this case, induced innovation—which functions quite differently from substitution—is the mechanism at play. A rise in the cost of carbon energy will encourage businesses to create new procedures and goods that use less carbon than their current offerings. As a result, they make unofficial investments in research, development, and new information. New items and processes are created by new knowledge, which reduces the output's *carbon intensity*. We have eliminated the processes of substitution from the R&DICE model to streamline the analysis and concentrate it on the induced-innovation mechanism. the emissions equation in the R&DICE model reflects the presumption that all increases in *carbon intensity* happen because of induced innovation rather than substitution. While in the common DICE model *carbon emissions* are depending on substitution effect, in the R&DICE, industrial *carbon emissions* depend only on output and endogenous *carbon intensity* σ(t).

$$E(t) = \sigma(t)Q^l(t) + E^l(t) \qquad \text{(1-R\&DICE)}$$

The DICE model undergoes one final modification when calculating *carbon intensity*. Estimates of technology developments are plugged in the common model to compute the no-controls *carbon intensity*, or the *carbon intensity* without any climate change policy in the form of regulations or carbon taxes. the *carbon intensity* in the R&DICE model is endogenous and is governed by the equation for induced innovation. According to this specification, *carbon intensity* and its change is a function of the amount of money spent on research for technical advancement that reduces carbon emissions.

$$[d\sigma(t)/dt]/\sigma(t) = \Psi^1 R(t)^{\Psi^2} - \Psi^3 \qquad \text{(2-R\&DICE)}$$

R(t) represents the energy sector's Research and Development (R&D) inputs for carbon-based energy (in induced-innovation approach); Ψ1 is the productivity of research; Ψ2 is the elasticity of technology to research; Ψ3 is the depreciation rate of technology. Given the specification, industry, and time frame, the elasticity is thought to range between 0.05 to 0.20. Given the technology and the time frame. the depreciation rate is variously estimated between 1 and 10 percent every year. These estimates are used as broad limitations to guarantee that the calibration generates accurate results. the calibrated equation used in Nordhaus (2014) research is equal to

$$\frac{[\sigma(t) - \sigma(t-1)]}{\sigma(t-1)} = 0{,}415\, R(t)^{0{,}139} - 0{,}20.$$

REMIND

The Regional Model of Investment and Development (REMIND), presented by Luderer (2013), is another IAM studying climate change that, differently from other IAM, models carbon emissions as a function of energy demand, types and sources of energy, and their carbon intensity, computed as a ratio between grams of CO2e and Kilowatt-hour (kWh) rather than tonnes of CO2e and million euros of revenues. the energy composition, source, technology of conversion into energy, and kind of energy are used to forecast the energy demand and the carbon emission for each place and period.

There are three distinct techniques to model *carbon emissions*, each dependent on different sources of energy. Given the types and sources of emissions, REMIND accounts for emissions at various levels of detail. According to sources, it determines the CO2 emissions from fuel burning, the Methane (CH4) emissions generated by the extraction of fossil fuels and domestic energy usage, and the Nitrous Oxide (N2O) emissions generated by the provision of energy. the energy system includes data on regional fossil fuel and biomass usage for each era and technological advancement. REMIND utilizes particular emissions factors, published by the EDGAR, that are calibrated to match base year GHG inventories for each fuel, area, and technology. Methods for reducing CH4, N2O, and CO2 from land-use change are not dependent on energy use. Three different approaches are used to establish baseline emissions, by source (and thus dependent on the emissions factors), by econometric estimate, or exogenously. the CO2 emissions from cement manufacture and the CH4 and N2O emissions from waste disposal are estimated using econometric tools. Either the growth of the Growth Domestic Product (GDP), which serves as a proxy for waste generation, or the capital investment determines the driver of emissions in each situation, as a proxy for cement production in infrastructure. For industry and transportation N2O emissions, REMIND uses exogenous variables retrieved from outsourced scenario data.

DEMETER

The Decarbonisation Model for Endogenous Technologies Emission Reductions (DEMETER), presented by B.C.C. van der Zwaana *et al.* (2015), takes into account different costs of carbon and non-carbon energy. Carbon energy technologies release GHG in the atmosphere during the production, such as fossil fuels. Non-carbon energy do not release any GHG. This type of technology includes wind and solar power. the model accounts for different technological level, technological improvement, and thus cost efficiency. Technological improvements are endogenous and depend on learning-by-doing, meaning that corporates become more efficient as they consumed more a certain type of energy. This efficiency is reflected in a lower cost for the same amount of energy produced. DEMETER assumes carbon intensity time-dependent and an empirical reduction rate, as in the DICE model. the use of carbon energy, is instead modelled as a function of the price of the two types of energy and the elasticity between the two. the price of non-carbon energy is reduced faster by the learning-by-doing effect since it is a rather new technology, still developing, compared to carbon energy. A scaling function g(), which is dependent on the cumulative capacity or cumulated energy delivered, incorporates learning-by-doing into the model. This scaling function illustrates that in comparison to when a high level of cumulative capacity is available, producing a given level of energy requires considerably more capital and Maintenance and Operations (M&O) costs when there is little cumulative capacity deployed.

The problem with this approach is that individual firms alone are not able, often, to internalize learning effects in their prices. Corporates can, anyway, learn from their peers. For the model that will be presented, the sector's cumulated capacity will be considered to have learning effects internalized by individual companies. the company's emissions will be a function of carbon intensity and activity. *Carbon intensity* is built for each industry, taking into account the *carbon emissions* that vary only due to changes in the size of the sales. the yearly change in the engineered *carbon intensity* is considered endogenous, and dependent on industry's R&D investments, cost of capital, and a time varying intercept. the size of the activity is considered to be endogenous as well and given the cost of capital and GDP as a proxy for energy demand.

Chapter 3

Literature Review

The impact of Finance

To study whether *carbon emissions* influence the riskiness of a company and its cost of capital, previous literature on climate finance has been considered. Climate finance is the branch focused on the relationship between investments and the impacts of climate change or the reduction or limitation of greenhouse gas emissions. In this branch, literature is split between two general streams: the first stream focuses on impact investors and the outcome from having responsible investment preferences. To this first stream are applied equilibrium models; the second focuses on Environmental, Social, and Governance (ESG) literature, and uses most modern asset pricing techniques theoretically and empirically, but it takes ESG scores as given and studies whether ESG scores, and not *carbon emissions*, have an impact. Results from both streams generally lead to controversial results.

In investing for impact in general equilibrium (Harris, 2022), investors with heterogeneous philanthropic preferences maximize their utility, which is a function of wealth, environmental impact, and impression of environmental impact, in an equilibrium model. This model considers both primary and secondary markets, and firms with heterogeneous impacts and endogenous size. It is crucial that firms have an endogenous size because the main channel through which investors have an impact in this model is the change of size. It's also possible for firms to change behaviours, but it is really fundamental for firms to change size depending on investors reallocating their capital. the research concludes that that is only with pure altruism, and not with the urgency of looking altruist, that investors can choose to put most of their money in firms reducing negative social impacts. It does not mean that holdings based on this urgency have no impact. Urgency holdings rely on influencing other investors in the primary markets to respond to its preferences, because the warm glow is focused on the secondary market, so, in the author's opinion, it can't proactively mandate more capital and put it in firms with high positive social impacts.

Climate Impact Investing (T. De Angelis *et al.* 2021) is another equilibrium model that studies the relationship between climate investments, which are investments made to reduce the impacts on climate changes, and their impact. the equilibrium model accounts for the percentage of climate investors on stock exchanges and the *carbon emissions* plans of the companies. the model seeks to study the effect of climate investors' negative screenings[4] and climate uncertainty[5] on invested companies' *carbon intensity*. Particularly, they discover that climate externalities (positive or negative) raise awareness in climate investors, which, through their screening, will contribute to influencing the cost of capital of negatively screened companies. When cost of capital is increased, a company is sometimes incentivized enough to invest in less carbon-intensive projects, to lower its cost of capital. When this happens, the *carbon intensity* of the company is decreased. Inversely, increasing uncertainty about the climate future, in terms of externalities such as climate change physical risk, regulation and technological innovation, tends to reduce companies' incentive to mitigate their emissions. Subsequently to climate uncertainty, *carbon intensity* increases.

The model presented in this research aims to study the same effect of concerned investors on cost of capital of carbon-intensive corporates, relying on empirical econometric analysis, and without making a distinction between environmentally concerned and non-environmentally concerned investors. We can observe that two general equilibrium models have controversial results on the role of secondary markets on the influence of corporates' social impact.

The impact of impact investing (Berk, and van Binsbergen, 2021) contributes to the literature of asset pricing models by studying the effect of ESG financial products. In the research, the quantitative impact of ESG divestitures is evaluated. For divestitures to have impact they must change the cost of capital of affected firms. Divestments must alter the cost of capital of companies in order to have an impact. the proportion of capital invested by socially conscious funds, the proportion of targeted firms in the economy, and the correlation between the targeted firms and the rest of the stock market are the three inputs that are used to derive a simple expression for the change in the cost of capital as a function of each. the authors show that the change in the cost of capital can be closely represented by a straightforward formula given the common assumptions that underlie the Capital Asset Pricing Model (CAPM), the basic model in financial economics. A divestment strategy will result in a change in the cost of capital that has a size depending positively on socially conscious investors group of target

[4] Negative screenings refer to the choice of ESG investor to not invest and to divest in companies harmful for the environment, the society, or their own corporate governance

[5] Climate uncertainty refers to the uncertain effects of climate change on natural disasters

companies as a percent of the economy, negatively on the correlation between these companies and the entire stock market, and positively on the historical market risk premium.

Results of this research show evidence of ESG divestures not being able to affect materially the cost of capital. This happens for the high correlation of ESG-targeted stocks with the rest of the market as well as for the small size of the wealth of socially conscious investors compared to the rest of the wealth on capital markets.

Differently from this study, the model that is going to be presented will not consider divestitures of socially concerned investors but will consider *carbon emissions* as a factor influencing the cost of capital of a company. Thus, the research considers *carbon emissions* rather than the fact of being in or outside a set of ESG target companies.

Derwall *et al.* (2011) point out that profit-seeking investors who think there is a temporary upward shift in predicted cash flows choose green companies. Most impressively, by using the screening strategy, investors can more effectively reach their ESG aim, but at a high cost: the cost of investing into an undiversified portfolio.

Instead, a sizable portion of socially conscientious investors pressure polluting corporations to make reforms, according to Heinkel et al. (2001). This study actually showed that the capital costs of socially irresponsible enterprises were higher than those of socially responsible businesses, which are excluded from the universe of eligible investments by investors. In his conclusion, he stated that an ESG factor will have a negative alpha if it holds a long position in green stocks and a short one in non-green stocks.

Pastor et al. (2021) also deny the existence of an ESG premium, and their study shows that this negative alpha is most apparent when financial markets' ESG preferences are stable.

The previously mentioned insight can be found in Pastor et al(2021) .'s empirical study of financial market equilibrium with green investors. Companies having a high overall ESG rating should, in equilibrium, have a negative alpha. the two components of a two-factor model are the ESG-specific and market factors. They also show how to develop the ESG factor, whose portfolio weights are determined by the target stocks' ESG ratings. the strategy used in this work for the production of the ESG factor is different from that used in Fama and French's (1992) methodology, which first identifies companies with high and low exposures to a business feature of interest before creating a long short portfolio.

Indeed, it is common to employ risk models that differ from the Fama and French factor model. Instead of developing a long-short portfolio of equities that are sorted by this attribute, Bolton and Kacperczyk (2021) create a carbon factor whose portfolio weights are proportional to enterprises' carbon emissions. Both utilising ESG ratings as a cardinal variable and using ESG ratings as an ordinal variable can be used to generate an ESG factor. This has been proven throughout time by empirical investigation. Using the ordinal variable technique, ESG ratings are utilised to determine a group of green and non-green stocks. Value-weighting is a common practise for stock returns within each cluster, and the return of the ESG is the difference between the returns of the green and non-green portfolios.

As stated by J.B. Berk, J.H. van Binsbergen (2021), risk models should beware of the fact that companies considered to have a negative societal impact appertain to few industries that may have specific risks for which the stocks have on average higher excess returns. the same probably happens when considering CO2e emissions, which are mainly concentrated in the energy and extractive industry. For this reason, the presented model will contain explanatory variables that control for the fact of belonging to these industries. the model presented in this research will exploit the risk factor approach, isolating the belongingness to a specific industry from the level of carbon emissions.

Chapter 4

Corporate carbon emissions

Corporate carbon emissions are greenhouse gas emissions that are caused directly or indirectly by business operations. These emissions can originate from a number of processes, such as the creation and consumption of energy, transportation, and the manufacture and consumption of goods. *Corporate carbon emissions* can be measured under three different Scopes, each in CO2e units. Direct emitting sources that are owned or controlled of an organisation produce scope 1 emissions. For instance, scope 1 covers the emissions from a trucking company's fleet of internal combustion engines. Consuming purchased power, steam, or other energy sources upstream from a company's primary operations results in scope 2 emissions. Scope 3 encompasses all other emissions associated with a company’s operations that are not directly owned or controlled by the company (S&P Global, 2022). Therefore, scope 3 emissions include several sources of indirect emissions in both the company’s supply chain and downstream from the company’s owned or controlled operations (e.g., the emissions from the in-use phase of a company’s products or services, such as the driving of a truck produced by an automobile manufacturer).

The most carbon-intensive companies appertain to few industries (Berk and van Binsbergen, 2021). Because the *carbon emissions* of a given business depend on a number of variables, including the processes and technologies employed, the industry's location, and the specific goods or services it provides, it is challenging to rank industries by *carbon intensity*. Additionally, when regulations and technological advancements change over time, an industry's *carbon intensity* may alter considerably. the following are 8 sectors that are usually regarded as being the most carbon-intensive:

Industry	Description
Energy production	The energy sector is a major contributor to global carbon emissions, with fossil fuel-based electricity generation being a particularly carbon-intensive activity.
Heavy manufacturing	Industries such as steel, cement, and aluminum production require significant amounts of energy to produce, leading to high carbon emissions.
Transportation	The transportation sector is a significant contributor to carbon emissions, with the burning of fossil fuels for transportation being a major source of CO2 emissions.
Agriculture	Agricultural activities such as livestock production and fertilizer use can also contribute to carbon emissions.
Chemical and petrochemical industries	These industries rely on fossil fuels for the production of chemicals, plastics, and other products, leading to high carbon emissions.
Mining	The extraction and processing of minerals and other raw materials can be energy-intensive, leading to high carbon emissions.
Refining and petrochemical processing	The refining and processing of oil and gas is a carbon-intensive activity.
Construction	The construction industry consumes large amounts of energy, particularly for the production of cement, steel, and other building materials.

Since the information on the three Scopes is reported after the end of the financial year, the published information starts to influence investors starting from the next financial period. That is why the past year's *carbon emissions* in the three different scopes are used in the presented model as explanatory variables of the current year's return.

Under the efficient market theory (Fama, 1962), investors should be able to price all company information that is available, thus, even *corporate carbon emissions*. These actors should all share the same information regarding GHG that are being emitted by past and future companies' projects and price this information with the same model. the problem with the current year's carbon emissions is the fact that they suffer from reverse causality with the company's stock return. In fact, *carbon emissions* and their increase are influenced by the size of the revenue and the increase in them, because higher revenues normally mean a larger company's activity which in turn means higher *carbon emissions* if that activity is a polluting one. As revenues increase *carbon emissions*, they also increase returns as they signal company growth and mean higher cashflows if the company's marginal profitability stays still. Is thus reasonable to think that higher revenues generate both higher *carbon emissions* and returns. Investors concerned with the risks coming from emitting GHG, *carbon emissions* price negatively the company, making returns lower. To overcome this issue of reverse causality, and still be able to include the effect that the current level of *carbon emissions* can have on the cost of capital, the TSLS estimation is used with instrumental variables. This estimation method allows to avoid the effect of reverse causality.

Chapter 5

Empirical Analysis

Cost of capital model

To understand how *carbon emissions* price stocks and influence their cost of capital, an extension of the Fama and French (1992) three-factors model is used. In addition to the three factors, scope 1, scope 2, and scope 3 absoulte emissions are added to the model together with a dummy variable that signals whether the company appertains to a carbon-intensive industry. This last variable is used in order to isolate industry-specific risks that are not linked to *carbon emissions*. Companies expected excess returns are explained by the following stochastic equation:

$$r_i = \beta_i{}' FF + \gamma_i E_i + \delta_i D_i + \varepsilon_i \quad (1)$$

r_i is the vector of excess returns of the stocks belonging to the considered market index

β_i is the vector of slope coefficients of the excess returns of asset *i* on the 3 factors excess returns (Fama and French 1992) and an intercept, and FF_t is the matrix made of an intercept and the vectors of the three factors excess returns.

$$\beta_i = [\beta^0, \quad \beta^1, \quad \beta^2, \quad \beta^3];$$
$$FF_t = [1, \quad MKT_t, \quad SMB_t, \quad HML_t];$$

MKT_t is the excess return of the market in year t. This value is the value weighted average excess return of the index.

SMB_t is the excess return made by the value weighted portfolio long on the first decile of stocks ordered in decreasing market capitalization size minus the last decile.

HML_t is the excess return made by the value weighted portfolio long on the first decile of stocks ordered in decreasing value of book to market ratio minus the last decile.

γ_i is the slope coefficient of the estimation of the excess returns of asset *i* on the sum of the three scopes of *carbon emissions*, and $E_{i,t}$ is the time series of the sum of the three scopes of carbon emissions.

$$E_{i,t} = \left[Scope^1_{i,t} + \quad Scope^2_{i,t} + \quad Scope^3_{i,t}\right];$$

$Scope^1_{i,t}$ are the tonnes of CO2e Scope 1 emissions of company i in year t.

$Scope_{i,t}^{2}$ are the tonnes of CO2e Scope 2 emissions of company i in year t.

$Scope_{i,t}^{3}$ are the tonnes of CO2e Scope 3 emissions of company i in year t.

δ_i is the slope coefficients of the excess returns of asset i on the intensive industry dummy variable, and $D_i$is the vector made of the carbon intensive industry dummy variable. the industry dummy variable is equal to 1 if the company appertains to carbon intensive industries, otherwise it is 0. the carbon intensive industries are the 8 reported in Table 1.

$D = [D^{ENG} + \quad D^{HVM} + \quad D^{TRN} + \quad D^{AGR} + \quad D^{CHE} + \quad D^{MIN} + \quad D^{REF} + \quad D^{CON}]$

D^{ENG} is a vector made of many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Energy Production industry, otherwise it is equal to 0.

D^{HVM} is a vector made of many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Heavy Manufacturing industry, otherwise it is equal to 0.

D^{TR} is a vector made of many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Transportation industry, otherwise it is equal to 0.

D^{AGR} is a vector made of many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Agricultural industry, otherwise it is equal to 0.

D^{HVM} is a vector made of as many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Heavy Manufacturing industry, otherwise it is equal to 0.

D^{CHE} is a vector made of as many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Chemical and Petrochemical industry, otherwise it is equal to 0.

D^{MIN} is a vector made of as many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Mining industry, otherwise it is equal to 0.

D^{REF} is a vector made of as many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Refining and Petrochemical processing industry, otherwise it is equal to 0.

D^{CON} is a vector made of as many elements as the number of companies. Each element is equal to 1 if the corresponding company belongs to the Construction industry, otherwise it is equal to 0.

Unique Estimation

Each coefficient in (1) is first uniquely estimated for every company and time period in the study through a Two Stage Cross Sectional (TSCS) regression. the TSCS regression is computed by averaging each variable in each period of time, and by regressing the time-varying average of the dependent variable on the time-varying average of the independent variables. Before averaging each variable, a time-series regression is done on FF_t, the 3 French and Fama factors with an intercept for each company.

In this way, it is possible to compute the coefficients over the three factors for each company. These coefficients are multiplied to the French and Fama factors for each year before being averaged. the estimation model used is an Ordinary Least Squares (OLS) regression, and is performed on the following expression:

$$E_t[r_i] = \beta E_t\left[\beta_i^{c\prime} FF\right] + \gamma E_t[E_i] + \delta E_t[D_i] \quad (2)$$

Time-varying estimation

Each coefficient in (1), is then computed uniquely for every company, but once for each time period, in order to observe how coefficients change over time. To estimate time varying coefficients, a Cross Sectional regression is performed in each period.

Before regressing the whole dataset in each period, a time-series regression is performed on X_t, the 3 French and Fama factors with an intercept for each company.

In this way, it is possible to compute the coefficients over the three factors for each company. Then, in each Cross-sectional regression, French and Fama Factors are multiplied by the company's coefficients in each observation.

The estimation model used is the Ordinary Least Squares (OLS) regression, and it is performed on the following expression:

$$r_i = \beta\left[\beta_i^{c\prime} X\right] + \gamma E_i + \delta D_i + \varepsilon_i \quad (3)$$

The cross-sectional estimation is repeated in each period. Thanks to the Cross-sectional estimation, it is possible to observe whether estimated coefficients vary over time in value and statistical significance.

The reverse causality of Greenhouse Gas emissions and returns

The expression (1), if estimated through a linear regression, would suffer of reverse causality. This happens because, when a company grows, its revenues increase together with its *carbon emissions*. When its revenues grow, the excess returns of a company generally increase, although the cost of capital may be influenced negatively by the increased *carbon emissions*. Reverse causality is a source of endogeneity, meaning that the error term will be correlated with the explanatory variable, and will thus make a linear regression model biased.

To avoid reverse causality, the model relies on a Two Stage Least Square Estimation (TSLS) estimation made with an instrumental variable (James and Singh, 1978). In a stochastic function, an instrumental variable (IV) is a variable that has a correlation with the relevant explanatory variable but is not correlated to the error term. In other words, the explanatory variable is "instrumented" using an IV to obtain accurate estimates of its effects on the response variable.

In this model, the endogenous explanatory variable is the sum of the 3 Scopes of absolute *carbon emissions* since they are correlated with the company's growth and thus its returns. the instrumental variables will be the 3 Scopes of *carbon intensity*. *Carbon intensity* is the ratio between corporate's *carbon emissions* and revenues. It can be computed for each of the three scopes. the *carbon intensity* of Scope *s* and company *i* is equal to

$$CI_i^s = \frac{Scope_i^s}{Revenues_i} \quad for\ s = [1, 2, 3]$$

Carbon intensity, in fact, must be highly correlated with *carbon emissions*, but since it does not take into account the growth of the company, and thus will not be influenced by growing revenues and returns. W*i* is the vector of company's *i* sum of 3 Scopes of *Carbon intensity*.

$$W_i = [CI_i^1 + CI_i^2 + CI_i^3]$$

In this model the vector of coefficients resulting from the Cross-Sectional TSLS regression is equal to the following variable β_t^{TSLS}

$$\beta_t^{TSLS} = \left(X_t{}' P_{z,t} X_t\right)^{-1} X_t{}' P_{z,t} Y_t$$

Where P_z and X are defined as follows

$$P_{z,t} = (Z_t{}' Z_t)^{-1} Z_t{}' Z_t$$

$$Z_t = \beta^{c\prime} FF_t + W_t + D_t$$

$$X_t = \beta^{c\prime} FF_t + E_t + D_t$$

As in the OLS estimation, the TSLS cross-sectional estimation is repeated in each period. This shows how the coefficients vary over time in value and statistical significance.

The TSLS estimation with IV is applied to the TSCS model as well. In this case the endogenous variable is the average of the sum of the three scopes CO_2e emissions, while the IV is the average of the sum of the three scopes *carbon intensities*.

Testing the Instrumental Variable

There are two requirements to be an acceptable IV. the first is to be correlated with the endogenous variable, the second is to be uncorrelated with the error of the explanatory equation of the model.

The Hausmann Test

A statistical technique called the Hausman test is used to decide whether to adopt a data model of panels with fixed effects or random effects. the test compares the variation in the coefficient estimates between the two models. It is used to decide if the unobserved effects are correlated with the model's independent variables, in which case a fixed effects model would be ideal, or whether they are not correlated, in which case a random effects model would be more suitable.

The test is based on the variance in the coefficient estimates between the two models, and the variance is assessed to see if it is statistically significant by comparing it to a chi-squared distribution. It is concluded that the two models are not different and that either one can be utilised if the difference is not statistically significant (Hausmann, 1981).

In econometrics and other subjects that employ panel data, the Hausman test is frequently employed. It can help to ensure that the estimates of the coefficients are accurate and efficient, making it an effective tool for choosing the right model to apply when evaluating panel data.

The Hausman test can also be used to gauge an instrumental variable's effectiveness in an economic model. the main benefit of using an IV is that it makes it possible to determine the causal connection between an independent variable and a dependent variable even when the model incorporates endogeneity (i.e., a correlation between the independent variable and the error term). This is due to the fact that the IV should be associated with the independent variable of interest rather than the error term in order to assist eliminate the bias caused by endogeneity.

The Hausman test can be used to determine whether the IV is a reliable instrument by comparing the estimates of the coefficients produced from an IV model with a reduced form model. When estimating the reduced form model, the dependent variable used is the endogenous variable, which is a function of the employed IV. the IV model uses this Instrumental Variable as the independent variable, which determines partially the endogenous ones, and other control variables.

The IV is probably a good instrument if there is a statistically significant difference between the coefficients from the two models because it can account for the endogenous variable's change.

To determine if the IV is a useful instrument, the Hausman test compares the estimates of the coefficients obtained from an IV model with a reduced form model, with statistically significant difference indicating a strong IV and small difference indicating a weak IV.

$$Hausmann\ Test\ Statistic$$
$$= (\beta^{TSLS} - \beta^{OLS})'[Var(\beta^{TSLS}) - Var(\beta^{OLS})]^{-1} (\beta^{TSLS} - \beta^{OLS})$$

Where $Var(\beta^{TSLS})$ and $Var(\beta^{OLS})$ are the corresponding variance-covariance matrices of the TSLS and OLS estimates respectively.

The chi-squared distribution of the test statistic has the same number of degrees of freedom as the coefficients in the model. If the test statistic overcomes the crucial value from the chi-squared distribution, the null hypothesis that the coefficients in the IV model are equal to the ones in the reduced form model is rejected, and the goodness of fit of the IV is proven.

The Hausman test's crucial value for the strength of an instrumental variable would be provided by the analysis's p-value. A p-value is a probability that expresses the level of significance of a test statistic. If the null hypothesis is true, there is a chance that the test statistic will be just as extreme as the one derived from the data, if not more so. In other words, if the null hypothesis is true, it is not possible to state with statistical confidence that the estimates made employing the IV fit the observations of the dependent variable better than the estimation made without using it.

The p-value is considered statistically significant and the null hypothesis is rejected if it is less than 0.05, which is a common threshold.

It is significant to remember that the p-value of 0.05 serves as the cutoff for deciding whether the null hypothesis should be accepted or rejected in many scientific fields. the researcher must decide what level of significance is appropriate for their topic because this value is arbitrary.

The Cragg-Donald Test

The Cragg-Donald test is a statistical procedure used to assess the reliability of an instrumental variable (IV). the Hausman test, which compares estimates of the coefficients from an IV model with a reduced form model, is a development of this test.

When doing the Cragg-Donald test, the dependent variable used is the endogenous variable, while the explanatory variable employed is the IV. Then, to perform the IV model's estimation, the endogenous variable is employed as the independent variable, and a set of control variables are utilised (Cragg and Donald, 1993).

The Cragg-Donald test statistic is determined as follows:

$$Cragg - Donald\ Test\ Statistic\ = \frac{\sqrt{[Var(\beta^{TSLS}) + \ Var(\beta^{OLS})]}}{(\beta^{TSLS}\ -\ \beta^{OLS})}$$

The test statistic is followed by a typical normal distribution. If the test statistic is overcomes the threshold of the critical value from the traditional normal distribution, the null hypothesis that the coefficients of the IV model are equal to the ones of the reduced form model is rejected, and the better fit of the IV is proven.

Since the Cragg-Donald test is less prone to over-identification bias, it is particularly useful when there are more instruments than there are participants in the study. the test is employed in situations when the instruments are subpar because it does not require as much exogeneity as the Hausman test.

The reliability of instrumental variables is examined using both the Cragg-Donald test and the Hausman test, although each test has its own set of assumptions and limitations. the researcher must decide which test best fits their study issue and data collecting.

The threshold for accepting estimates in both the Cragg-Donald and Hausman tests is determined by the p-value, a measurement of the test statistic's statistical significance. the null hypothesis would be disregarded if the p-value was less than 0.05, a conventional requirement for determining statistical significance.

The Cragg-Donald test's null hypothesis states that the coefficients resulting from the employment of the IV are equal to the ones obtained in the reduced form model, whereas the alternative is that they differ. the test statistic has a conventional normal distribution. the IV is a reliable tool if the p-value is less than 0.05, which indicates that there is a statistically significant difference between the estimations of the coefficients of the IV and the reduced form model.

It is significant to remember that the p-value of 0.05 serves as the cut-off to decide whether accepting or rejecting the null hypothesis in many scientific fields. the researcher must decide what level of significance is appropriate for their topic because this value is arbitrary.

It's crucial to remember that a p-value of less than 0.05 does not necessarily imply that a study's findings are accurate; rather, it just indicates that they are unlikely to be the result of chance. So, one of the requirements for proof is a p-value lower than 0.05, but other factors like external validation and replication should also be taken into account.

Data Calibration

This script is written in MATLAB and performs a data calibration on a dataset taken by Refinitiv over the time frame of 2012-2021. the dataset is made of the biggest publicly traded firms in Europe and the United States, respectively, is reflected in the STOXX 600 and the S&P 500 stock market indices. the S&P 500 is made up of 500 firms listed on American stock exchanges, whereas the STOXX 600 is made up of 600 companies listed on stock exchanges in the European Economic Area. It is possible to learn about the economic climate and market trends in Europe and the United States by observing the returns of the stocks included in these two indices. There are probably many missing numbers in the dataset of the STOXX 600 and S&P 500 firms' absolute *carbon emissions*, which can complicate analysis and interpretation. One significant problem is that missing data may impact the size of the dataset and its representativeness, which may result in skewed or erroneous results. Furthermore, missing values can cause issues with the necessary assumptions and computations, making it challenging to conduct some analyses or statistical tests. the dataset is been retrieved in excel and then loaded into MATLAB. the script loads the data from a .mat file that contains the data for the European and American stocks in the two respective indices. the data is then converted into arrays and stored in variables containing the name of the stock, revenues, returns, scope 1 CO2e emissions, scope 2 CO2e emissions, scope 3 CO2e emissions, the index weighted return, the High minus Low factor (Fama and French 1993), the Small minus Big factor (Fama and French 1993), and the industry of each company.

The script then creates new variables by performing calculations on the loaded data. the variables "CI1", "CI2", and "CI3" are created by calculating the *carbon intensity* of Scope 1, 2, and 3 respectively. These variables are obtained by dividing the three scope emissions by the revenues of the company in each period.

The script also creates a "dummy matrix" of intensive industries by creating a new array "I_D" of zeros and then looping through the "Industry" array to identify and mark the companies that belong to intensive industries listed in Chapter 4.

The script then loops through each stock, and uses a linear regression model for each company to compute the vector β_i. If the company does not have any data available, the script assigns zeros to the beta value for that company.

To overcome the fact that many data are missing in the dataset, a particular data calibration is performed before computing the OLS and TSLS Cross-Sectional estimates. This code is performing a loop over a range of periods, from 2012 to 2021, and within each iteration it is performing several steps to avoid missing data in the dataset and still being able to perform the two regressions.

The code avoids to consider missing values by using functions to identify any missing or infinite value in the returns, in the sum of the three scopes emissions, and in the sum of the three *carbon intensities*. Any row containing a missing or infinite value is then removed from all the variable of that company in that time frame.

The normalisation phase in the procedure is used to give the total of the scopes and carbon intensities with the same scale of variance of returns. Each value of the explanatory variable X is divided by the square root of the sum of the squared differences between X and the sample average of X to achieve this. In this way, each variable is standardised.

Standardizing a variable is a common data pre-processing step that makes sure all the variables are on the same scale and can be directly compared in a regression analysis. Without standardisation, variables that are evaluated on different scales (for instance, in different units) may have a disproportionate impact on the regression analysis due to their scale.

The variance of the explanatory variable and the instrumental variable are adjusted to have a mean of zero and the same variance of the excess returns by normalising the sum of the three scopes of carbon emissions and the three scopes of carbon intensity.

The use of standardised explanatory and instrumental variables in the regression would be incorrect since the research examines the impacts of carbon emissions and carbon intensities on excess returns throughout a range of time periods, and because the variance of excess returns varies over time. For this reason, the excess returns' standard deviation of each period is multiplied by the standardised sum of the three scopes of carbon emissions and carbon intensities for each time frame.

As a result, the influence of X and Z on the dependent variable can be compared because they are engineered to have the variance of the excess returns. This also makes it easier to interpret the regression coefficients because they will be interpreted as the variation of the dependent variable per unitary variation of the standardised explanatory variable.

Chapter 6

Empirical Results

TSCS Regression with Carbon emissions

Ordinary Least Squares Regression

The Coefficients resulting from the TSCS OLS regression containing the three French and Fama factors, the intensive industry dummy variable, the sum of the three scopes of absolute CO2e emissions, and an intercept are as follows:

TSCS OLS Regression Coefficients EU

Explanatory Variable	Coefficient	Value
Intercept	α	116.43
MKT Factor	β1	-0.07
SMB Factor	β2	-0.61
HML Factor	β3	0.11
Intensive Industry Dummy	δ	309.66
3 Scopes CO2e sum	γ	3.31

TSCS OLS Regression Coefficients US

Explanatory Variable	Coefficient	Value
Intercept	α	34.64
MKT Factor	β1	0.55
SMB Factor	β2	0.77
HML Factor	β3	0.00
Intensive Industry Dummy	δ	96.06
3 Scopes CO2e sum	γ	-0.95

We can already observe that the results of the regression suggest the non-predictivity of the model. In fact, in both cases α is very far from 0, and in the STOXX 600 dataset the sum of the three French and Fama factors coefficients is not close to 1.

Two Stages Least Squares Regression

The Coefficients resulting from the TSCS TSLS regression containing the three French and Fama factors, the intensive industry dummy variable, the sum of the three scopes of absolute CO2e emissions, the sum of the three scopes *carbon intensity*, and an intercept are the following:

TSCS TSLS Regression Coefficients EU

Explanatory Variable	Coefficient	Value
Intercept	α	5.92
MKT Factor	β1	0.06
SMB Factor	β2	-1.02
HML Factor	β3	-0.13
Intensive Industry Dummy	δ	-1211.26
3 Scopes CO2e sum	γ	2.36

TSCS TSLS Regression Coefficients US

Explanatory Variable	Coefficient	Value
Intercept	α	-20.77
MKT Factor	β1	-0.03
SMB Factor	β2	-2.93
HML Factor	β3	-0.46
Intensive Industry Dummy	δ	-1033.98
3 Scopes CO2e sum	γ	11.71

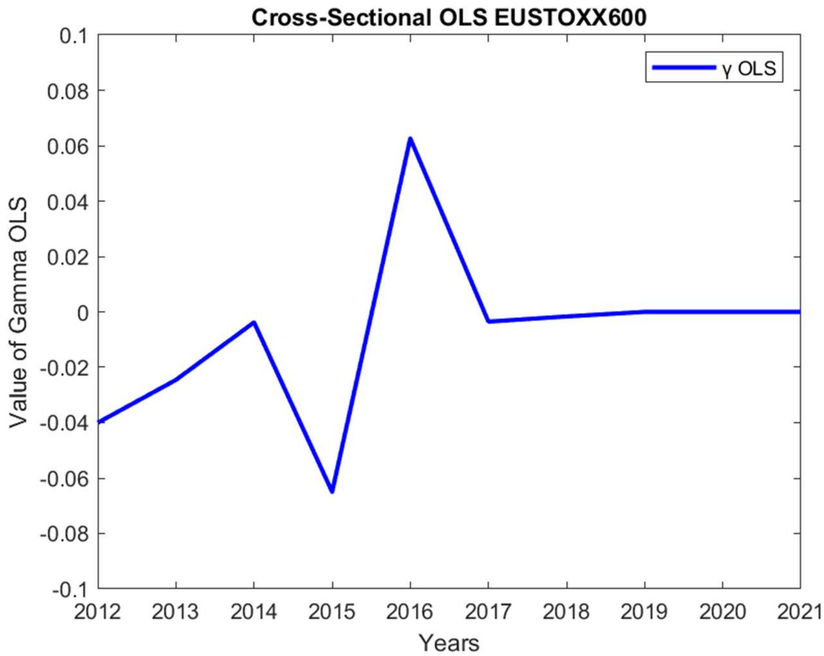
It is possible to observe from the results of the regression that the model foes not appear to be predictive. In fact, in both cases α is very far from 0, and the three French and Fama factors' coefficients is negative or close to 0. the value of the resulting coefficient of the intensive industry dummy variable is exceptionally high. That is why it is possible to wonder that the source of bias in the predictive regression is exactly this dummy variable.

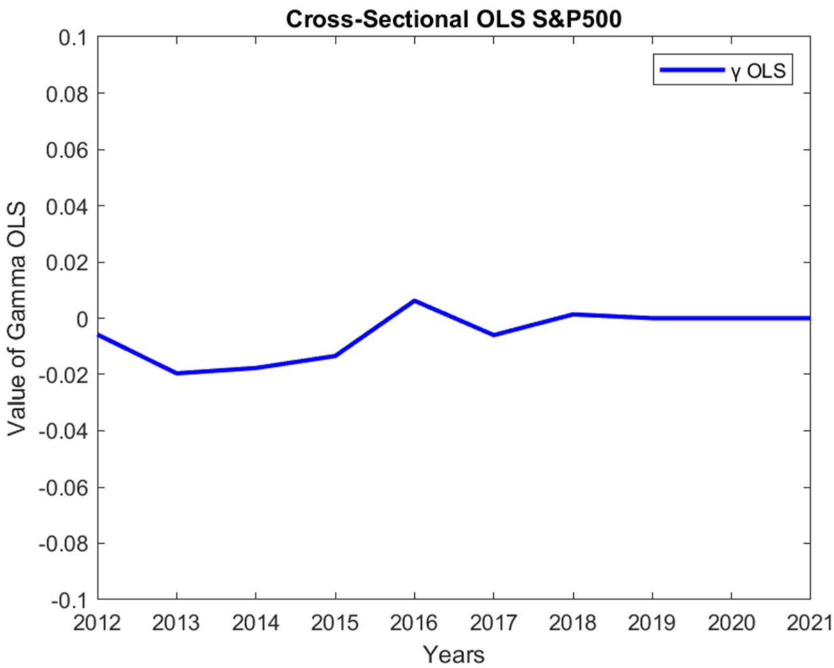
For this reason, the same estimation models should be performed on the same stochastic function, without the intensive industry dummy variable. Unhopefully, an estimation model without that variable may not account for a higher riskiness of carbon intensive industries that is not directly linked to *carbon emissions*.

Cross-Sectional Regression with Carbon emissions

Ordinary Least Squares Regression

The coefficients resulting from the Cross-Sectional OLS regression made in the time frame 2012-2021 are displayed in a 6 by 10 matrix, where there are 6 different coefficients in 10 observed periods, hence once per year. the objective of this study is to observe the behaviour of gamma (γ), the *Carbon emissions* coefficient, and its significance over time. the followings are the results of gamma between 2012 and 2021 for the STOXX 600 and the S&P 500 companies:



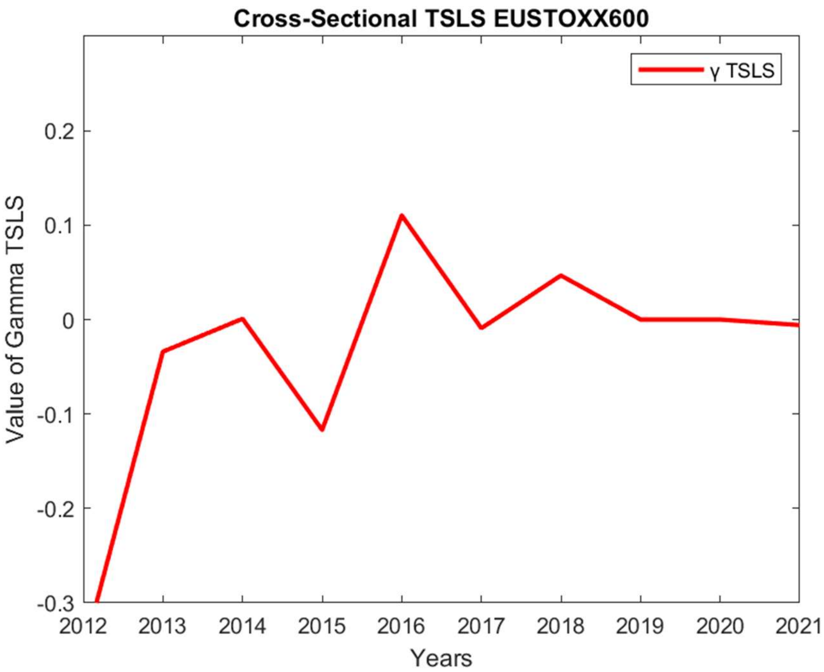


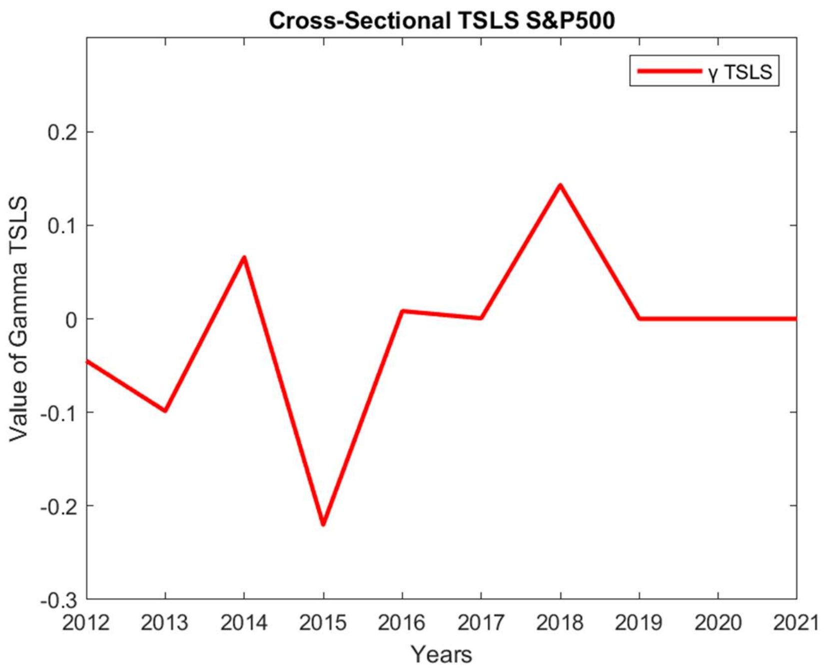
It is possible to observe that in both indexes coefficients are really flat after 2017. In both datasets, coefficients are negative except for the year 2016. When the gamma coefficient is negative, higher absolute *carbon emissions* bring to lower returns. In years when gamma is negative, the risk linked to high *carbon emissions* lowers the price of stock price of companies with high carbon emissions through investors' divestments.

In years where the coefficient is positive, companies with high *carbon emissions* are already undervalued because their risk is been accounted and thus they will overperform. This overperformance is not free of risk since those companies underperform in periods when investors tend to be more conscious of and adverse to risks linked with high *carbon emissions*. It is thus understandable that periods with negative coefficients anticipates periods with positive coefficients, just as in the estimated results. It is thus possible to observe that between 2012 and 2015, investors expectations regarding *carbon emissions* risk, made lower the price of companies with high CO2e absolute emission. This resulted in an overperformance of these companies in the next years, when this risk is not materializing. Gamma coefficients are between the bounds of -0,02 and 0,02 in the S&P 500 dataset, while they can variate between -0,08 and 0,08 in the STOXX 600 dataset.

Two Stages Least Squares Regression

The coefficients resulting from the Cross-Sectional TSLS regression made in the time frame 2012-2021 are displayed in a 6 by 10 matrix, where there are 6 different coefficients for 10different observed periods, hence once per year. the objective of this study is to observe the behaviour of gamma (γ), the coefficient of *Carbon emissions* projected on *Carbon intensities*, and its significance over time. the followings are the results of Gamma between 2012 and 2021 for the STOXX 600 and the S&P 500 companies:





It is possible to observe that in both indexes coefficients become flat after 2019. In the STOXX 600 dataset, coefficients are negative until year 2015. In 2016 and 2018 the coefficients are slightly positive, while in the remaining years the coefficients are zero or very close to zero. In the S&P 500 dataset, instead, coefficients are negative in 2012, 2013 and 2015, and positive in 2014 and 2018, being close to zero in the rest of the observations.

Differently from the coefficients resulting from the OLS estimations, the TSLS estimated coefficients vary between -0,3 and 0,2. Specially for the S&P 500 dataset, it is possible to observe that coefficients are not flat to 0 over time, when accounting for *carbon intensity* as Instrumental Variable.

Greater absolute *carbon emissions* projected on *carbon intensity* reduce returns when the gamma coefficient is negative. the risk associated with companies that report high *carbon emissions,* and divestitures made on them, affects the price of the stocks in years when gamma is negative.

Companies with large *carbon emissions* are already undervalued in years when the coefficient is positive because their risk has already been accounted, and thus their returns will outperform. Since these companies underperform during times when risks linked to high *carbon emissions* materialize, their overperformance is not risk-free, but is a compensation for risk.

Chapter 7

Robustness

TSCS Regression with Carbon emissions

Ordinary Least Squares Regression

TSCS OLS Regression Coefficients US

Explanatory Variable	Coefficient	P-value
3 Scopes CO2e sum	γ	0.9260

TSCS OLS Regression Coefficients US

Explanatory Variable	Coefficient	P-value
3 Scopes CO2e sum	γ	0.8828

Two Stages Least Squares Regression

TSCS TSLS Regression Coefficients EU

Explanatory Variable	Coefficient	P-value
3 Scopes CO2e sum	γ	0.9356

TSCS TSLS Regression Coefficients US

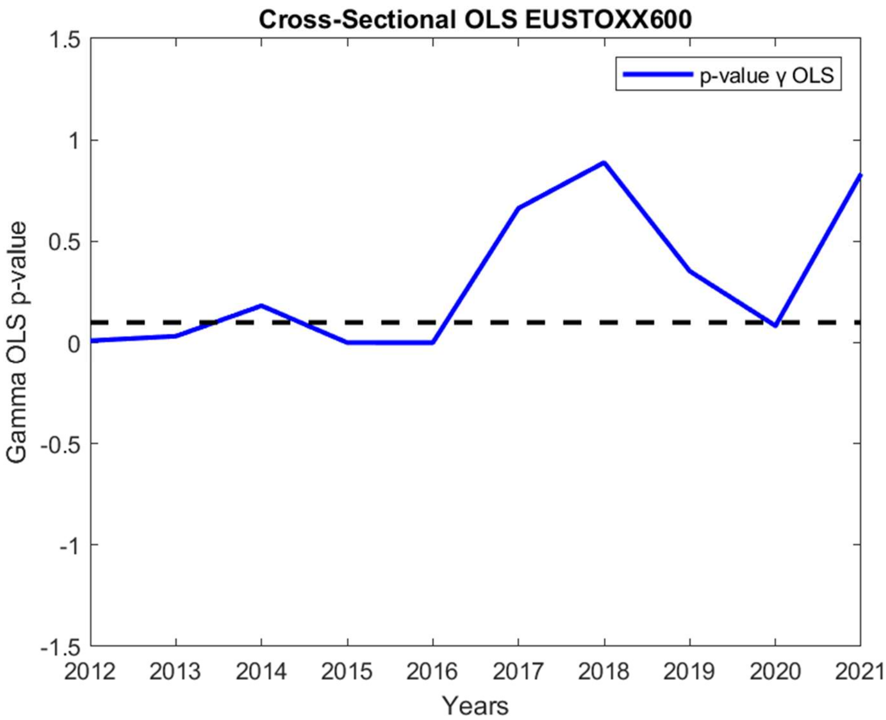
Explanatory Variable	Coefficient	P-value
3 Scopes CO2e sum	γ	0.8049

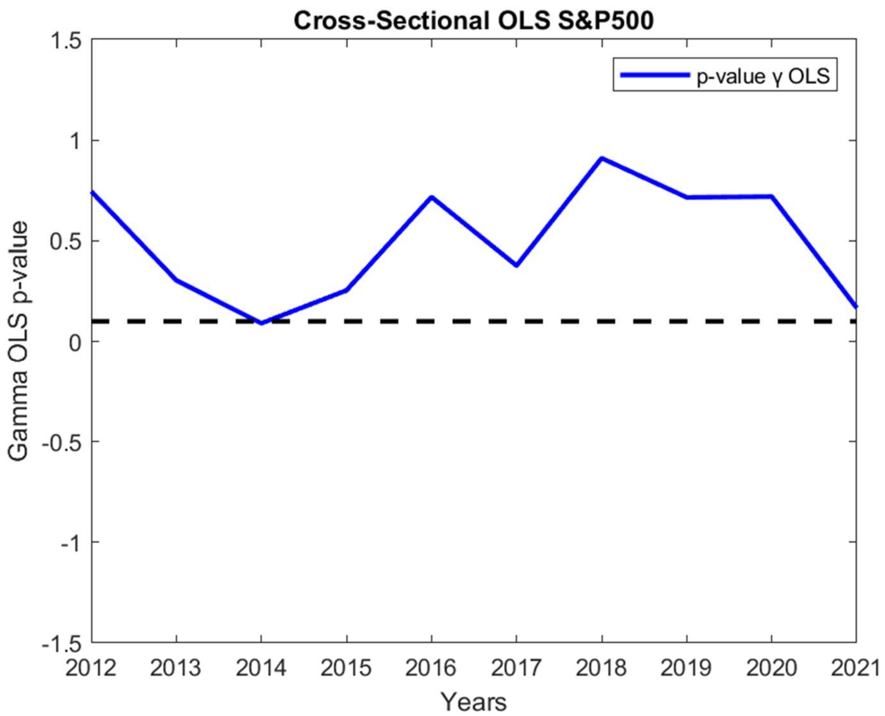
As it is possible to observe from the p-value of the TSCS Regressions performed on the two datasets, neither the OLS estimation nor the TSLS estimation are able to compute a carbon emission coefficient that is significantly different from zero.

Cross-Sectional Regression with Carbon emissions

Ordinary Least Squares Regression

These are the plots of the p-values of the Cross-sectional OLS regressions over the observed periods.





Is it possible to observe that, while in the STOXX 600 dataset there are multiples statistically significant observations of the coefficient gamma, in the S&P 500 dataset there is one statistically significant coefficient.

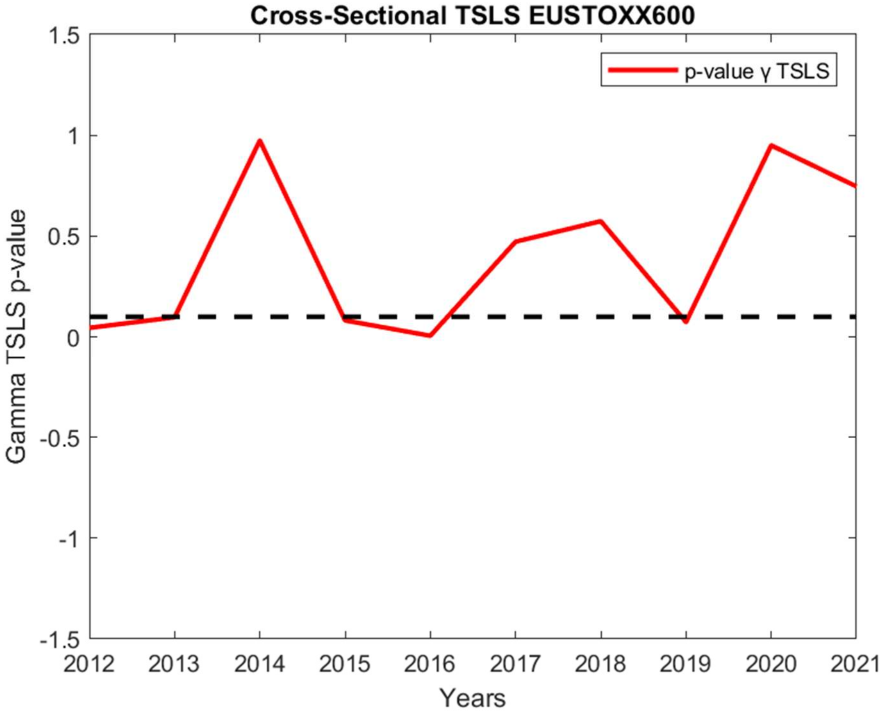
The statistically significant gamma coefficient of the S&P 500 dataset is observed in year 2014. the coefficient computed in this period is approximatively equal to -0,02.

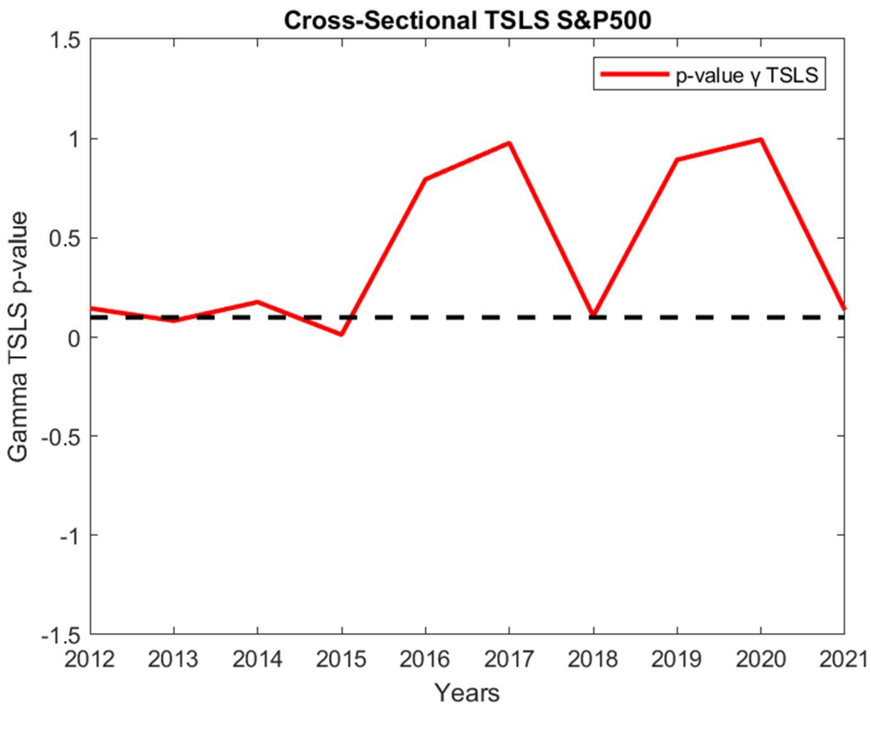
For the STOXX 600 dataset, instead, the statistically significant results are 5 out of 10. the first years in which the estimated gamma coefficient is statistically significant are 2012 and

2013. During these periods the coefficients are equal to -0,04 and -0,02 respectively. In 2014 the coefficient is not significant, and it comes back to be significant in 2015 and 2016, when its value passes from -0,06 to 0,06. the coefficient is then significant again in 2020, when the estimation is still very close to 0.

Two Stages Least Squares Regression

These are the plots of the p-values of the Cross-sectional TSLS regressions over the observed periods.





After plotting the p-value of the gamma coefficient estimation in the TSLS Cross-Sectional regression, it is possible to understand visually how the regression's fit improved, especially for in the S&P 500 dataset.

While in the OLS Cross-Sectional regression only one estimated coefficient is significant over ten observed periods, in the TSLS Cross-Sectional regression four observations out of ten are statistically different from 0.

In the estimation of gamma made with the TSLS regression on the STOXX 600 dataset, five estimations out of ten are significant. the estimated coefficients are significant in the years when they are significantly estimated in the OLS regression except for one period. In this model, the coefficient estimated in 2020 is not significant anymore, but the one estimated in 2019 is significant.

During the years 2012 and 2013, the STOXX 600 gamma coefficient is statistically significant and equal to -0,35 and -0,03 respectively. In years 2015 and 2016 the significant gamma passes from -0,12 to 0,11. In 2019 the coefficient is almost zero.

In 2013 the gamma estimated for S&P 500 companies is significant and equal to -0,1. the coefficient comes back to be significant in 2015 when its value is still negative and equal to -0,22. Then it is significant again in 2018 and 2021 when its value is equal to 0,14 and very close to 0 respectively.

Tests for the goodness of the Instrumental Variable

Hausmann Test

The Hausmann Test has a p-value that is below 0.05 for each observation both in the STOXX 600 and in the S&P 500 dataset. the test's null hypothesis is accepted in each period of the two observed dataset, and thus *carbon intensities* are valid Instrumental Variables.

Cragg-Donald Test

Even the test for weak Instrumental Variables, the Cragg-Donald Test, has a smaller p-value than the null hypothesis' threshold in each observation of the STOXX 600 and S&P 500 datasets. the null hypothesis of the test is rejected for the two observed datasets in each period observed. It is possible to state that, in this model, *carbon intensities* cannot be considered weak instrumental variables, because the coefficients estimated using *carbon intensities* as IV are statistically different from coefficients estimated with the OLS regression.

Chapter 8

Conclusions

In conclusion, the research aims to study the relationship between a company's *carbon emissions* level and its cost of capital. the research finds out that *carbon emissions* level has a significant and positive impact on the cost of capital of a company in some years. This result must be interpreted as a support to the idea that the stock market accounts for the negative externalities of *carbon emissions*, and incorporates them into the cost of capital of companies. the research also attempted to isolate the impact of *carbon emissions* on the cost of capital by controlling for other factors that may influence the cost of capital such as revenues, the Fama and French factors, and industry dummies.

To further analyze the robustness of the estimates, the research performed several regression models using different estimation techniques, including OLS and TSLS, on two datasets, the S&P 500 and STOXX 600. the results showed that the *carbon emissions* level had a significant positive impact on the cost of capital even after controlling for other factors in both datasets. However, the significance of the estimates varied between the two datasets and the different regression models used.

For example, in the OLS Cross-Sectional regression of the S&P 500 dataset, only one estimated coefficient was significant out of ten periods of observation. the TSLS Cross-Sectional regression performed on the same dataset showed four significant observations out of ten periods. In the STOXX 600 dataset, the significant observations remained relatively consistent between the OLS and TSLS models, with five significant observations out of ten periods in both models. the coefficient values for the significant observations also varied between the two datasets and the different regression models.

Additionally, the research performed a Hausmann test, which confirmed the validity of the instrumental variable employed in the TSLS regression. This supports the conclusion that the significant positive impact of *carbon emissions* on the cost of capital is robust and not due to any potential endogeneity biases or reverse causality. It also shows the evidence that the reverse causality between *carbon emissions* and returns exists, and using *carbon intensity* as Instrumental Variable is a possible way to avoid the estimation bias.

In summary, the findings of this research provide evidence that the stock market is accounting for the negative impact of *carbon emissions* on the cost of capital and that companies that emit higher levels of carbon have a higher cost of capital. This finding has important implications for companies, managers, and investors as they can use it to make more informed decisions about investing in low-carbon projects and reducing *carbon emissions*. Additionally, it highlights the need for policymakers to implement measures that encourage companies to reduce their *carbon emissions*, such as carbon taxes and regulations.

Chapter 9

Appendix

OLS Cross-Sectional Regression Coefficients – STOXX 600

Coefficient	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
α	23.88	32.88	6.24	8.27	11.86	23.25	-4.55	26.38	0.50	24.85
β1	0.00	-0.06	0.04	1.10	0.32	1.82	0.09	0.28	0.05	0.01
β2	0.03	0.18	0.48	0.82	0.54	0.65	0.23	0.30	-0.02	0.07
β3	0.04	0.12	0.90	-0.32	0.61	0.51	0.19	0.42	-0.04	0.12
δ	7.29	-9.78	-1.32	-2.88	-1.99	-6.78	9.71	-9.25	10.08	4.48
γ	-0.04	-0.02	0.00	-0.06	0.06	0.00	0.00	0.00	0.00	0.00

OLS Cross-Sectional Regression Coefficients – S&P 500

Coefficient	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
α	23.21	32.43	15.23	1.12	16.15	18.88	-1.36	33.36	12.53	30.59
β1	0.01	0.01	0.03	0.08	0.39	0.29	3.12	-0.09	0.00	-0.04
β2	0.45	0.81	0.19	-0.07	-0.35	1.58	-0.06	0.10	0.01	0.17
β3	0.68	0.52	0.15	0.15	0.60	0.94	0.08	0.33	2.05	0.04
δ	-7.29	0.73	9.66	-17.33	-0.90	-5.80	-4.54	-13.09	-16.19	1.68
γ	-0.01	-0.02	-0.02	-0.01	0.01	-0.01	0.00	0.00	0.00	0.00

TSLS Cross-Sectional Regression Coefficients – STOXX 600

Coefficient	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
α	23.82	32.86	6.02	8.26	11.74	23.24	-4.16	26.57	0.57	24.85
β1	0.00	-0.06	0.03	1.10	0.31	1.83	0.08	0.31	0.05	0.01
β2	0.02	0.18	0.50	0.82	0.55	0.66	0.24	0.33	-0.02	0.07
β3	0.04	0.12	0.93	-0.32	0.64	0.51	0.18	0.43	-0.04	0.12
δ	10.21	-9.14	-0.91	-2.87	-1.50	-6.12	7.19	-14.49	8.74	4.78
γ	-0.35	-0.03	0.00	-0.12	0.11	-0.01	0.05	0.00	0.00	-0.01

TSLS Cross-Sectional Regression Coefficients – S&P 500

Coefficient	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
α	22.80	32.19	16.72	-0.43	16.17	19.00	-0.81	33.37	12.57	30.70
β1	0.01	0.01	0.03	0.09	0.39	0.29	3.23	-0.09	0.00	-0.04
β2	0.45	0.79	0.20	-0.11	-0.35	1.57	-0.10	0.10	0.01	0.17
β3	0.69	0.51	0.20	0.13	0.60	0.93	0.07	0.33	2.05	0.04
δ	-4.77	3.86	1.46	-8.37	-1.13	-6.53	-7.72	-13.19	-16.60	0.62
γ	-0.04	-0.10	0.07	-0.22	0.01	0.00	0.14	0.00	0.00	0.00

TSLS Cross-Sectional Regression Hausmann Test p-vaules – STOXX 600

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Test p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TSLS Cross-Sectional Regression Hausmann Test p-vaules – S&P 500

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Test p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Bibliography

Berk, J.B., van Binsbergen, J.H. (2021). the impact of impact investing. George Mason University Law & Economics Research Paper Series, 21-26.

Bolton, P. and M. Kacperczyk (2021). Global pricing of carbon-transition risk. Technical report, National Bureau of Economic Research.

Cheema-Fox, A., B. R. LaPerla, G. Serafeim, D. Turkington, and H. S. Wang (2019). Decarbonization factors. Forthcoming Journal of Impact & ESG Investing, Special Fall Climate Issue.

Climate Model Intercomparison Project. (2021). CMIP6 Data Request. WCRP CMIP.

Collins, M., et al. (2012). the HadGEM2-ES Computational Climate Model. Journal of Advances in Modeling Earth Systems, vol. 4, no. 1, 94-116.

Cragg, J. M., and Donald, S. G. (1993). Instrumental variables estimation of nonlinear rational expectations models. Econometrica, 61(4), 953-966.

Derwall, J., K. Koedijk, and J. Ter Horst (2011). A tale of values-driven and profit-seeking social investors.

De Angelis, T., Tankov, P., Zerbib O. D. (2021). Climate Impact Investing. No 676, Carlo Alberto Notebooks, Collegio Carlo Alberto.

EDGAR. the Emissions Database for Global Atmospheric Research. Methodology (edgar.jrc.ec.europa.eu/methodology).

Fama, E. F. and K. R. French (1992). the cross-section of expected stock returns. the Journal of Finance 47 (2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics 33 (1), 3–56.

Harris J. (2021). Investing for Impact in General Equilibrium. Available at SSRN.

Hausman, J. A. (1981). Speculations on the Limits of Inference with Panel Data. Journal of Econometrics, 16(1), 51-74.

Heinkel, R., A. Kraus, and J. Zechner (2001). the effect of green investment on corporate behaviour. Journal of financial and quantitative analysis 36 (4), 431–449.

Hong, H. and M. Kacperczyk (2009). the price of sin: the effects of social norms on markets. Journal of financial economics 93 (1), 15–36.

Intergovernmental Panel on Climate Change. (2007). Climate Change 2007: the Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (Cambridge Univ Press, Cambridge, UK).

Intergovernmental Panel on Climate Change. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the IPCC (Cambridge University Press, Cambridge, UK).

Intergovernmental Panel on Climate Change. (2018). Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty (Cambridge University Press, Cambridge, UK).

Intergovernmental Panel on Climate Change (2022) Climate Change 2022: the Physical Science Basis, Contribution of Working Group II to the Sixth Assessment Report of the IPCC (Cambridge University Press, Cambridge, UK).

James, L. R., and Singh, B. K. (1978). An introduction to the logic, assumptions, and basic analytic procedures of two-stage least squares. Psychological Bulletin, 85(5), 1104–1122.

Jones, C. D., et al. (2010). the HadGEM2 Climate Model: Configuration and Evaluation." Journal of Advances in Modeling Earth Systems, vol. 2, no. 2, 127-144.

Luderer G, Leimbach M, Bauer N, et al (2013) Description of the REMIND model (version 1.6) SSRN Working Paper 2697070.

National Center for Atmospheric Research (2021). Community Earth System Model

Nordhaus, W. (2014). Modeling Induced Innovation in Climate-Change Policy. Technological Change and the Environment, 189–209.

Nordhaus, W. (2016). Projections and uncertainties about climate change in an era of minimal climate policies. Working Paper. National Bureau of Economic Research.

Nordhaus, W. (2018). Climate change: the Ultimate Challenge for Economics. Nobel Prize Lecture, December 8, 2018 by William D. Nordhaus Yale University, USA.

OECD: Organization for Economic Co-operation and Development. Data by theme; Environment. Greenhouse Gas emissions by source. (stats.oecd.org)

Pastor, L., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. Journal of Financial Economics 142 (2), 550–571.

Refinitiv (2022). Carbon dioxide equivalent emissions from Refinitv.

S&P Global (2022). Trucost ESG Analysis, Frequently Asked Questions, *What are GHG Protocol's scopes 1, 2, and 3?* (spglobal.com)

Schmidt, T. S. (2014, March 26). Low-carbon investment risks and de-risking. Nature.

Shaftel, H., Jackson, R., Callery, S., & Bailey, D. (2020). Climate change: How do we know? NASA Global Climate Change: Vital Signs of the Planet, Earth Science Communications Team, California Institute of Technology.

Van der Zwaana, B.C.C., R. Gerlagha, G. Klaassenc, and L. Schrattenholzer (2002). Endogenous technological change in climate change modelling. Energy Economics, Volume 24, Issue 1, 1-19.

Watanabe, M., et al. (2011). MIROC5: Model Description and Basic Results of CMIP5 20th Century Simulations. Geoscientific Model Development, vol. 4, no. 2, pp. 845-872.

Watanabe, M., et al. (2012). MIROC-ESM 2010: Model Description and Basic Results of CMIP5-20c3m and Historical Simulations. Geoscientific Model Development, vol. 5, no. 5, 961-986.

World Bank. (2021). Industry, value added (% of GDP). World Bank National Accounts Data, and OECD National Accounts Data files.