Università
Ca'Foscari
Venezia

Master's Degree programme

in Economics and Finance Second Cycle (D.M.270/2004)

Final Thesis

Default Probability Estimation for Green Mortgages: Empirical Evidence from Italy

Supervisor
Ch. Prof. Monica Billio

Graduand
Chiara Borghese
Matricolation number
868139

Academic Year
2018 / 2019

Contents

Introduction

Nowadays, more attention is put on the concept of green finance and evidences on that are the discussions in course on this theme.

The increasingly warning coming from environmental threats pushes more and more towards a sustainable development with financial institutions representing, due to their intermediary role, a fundamental point in promoting these initiatives in order to reach European Union predefined goals.

Nevertheless, these initiatives do not depend only on banks, but also from private investors, which have to enhance their interest on these themes.

This thesis, which focuses only on a portion of green finance, i.e. green mortgages, has the scope to highlight the importance on this theme and the advantage from this choice.

In particular, the goal is to demonstrate how the probability of default of borrowers decreases when we are dealing with energy efficient houses, in contrast with not efficient ones.

The idea behind is that more energy efficiency will turn out in lower utility bills, resulting in more monthly disposable income available to the borrower for repaying debt. At the same time the value of houses will increase, leading to lower loss for banks in case of default.

The work is subdivided in five chapters.

Chapter 1 tries to explain what green finance is through presentation of various definitions, while giving a general outlook of both the international and Italian situation. The role of bank in the development of this new branch of finance is also pointed out.

Chapter 2 provides firstly an overview of the housing situation in Europe and in Italy, and secondly an analysis about mortgages popularity. It presents also the Energy Efficient Mortgage Initiative.

Chapter 3 deals with the statistical methodologies employed for the study performed in this work.

Chapter 4 describes how the dataset was created and the choice of variables.

Finally, chapter 5 presents results obtained applying methods chosen and the interpretation of outcomes.

1. Green Finance

1.1 Definition

Nowadays, the concept of ‘green finance’ or ‘green financing’ is becoming more and more widespread even if there is no a clear and common way to define this new branch of finance.

The Green Finance Group, established in 2016 under the China’s G20, defines ‘green finance’ as≪ financing of investments that provide environmental benefits in the broader context of environmentally sustainable development. [⋯] Beyond the financing of green investments, green finance also involves efforts to internalize environmental externalities and adjust risk perceptions in order to boost environmentally friendly investments and reduce environmentally harmful ones [⋯] ≫[1].

In August 2016, China issued “Guidelines for Establishing the Green Financial System”, where defines ‘green finance’ as ≪financial services provided for economic activities that are supportive of environment improvement, climate change mitigation and more efficient resource utilization. These economic activities include the financing, operation and risk management for projects in areas such as

[1] Kahleborn, Walter, A. Cochu, I.Georgiev, F. Eisinger, D. Hogg, *Defining “green” in the context of green finance*, Luxembourg Pubblication Office of the European Union 2017, pag 1

environmental protection, energy savings, clean energy, green transportation, and green buildings≫[2].

Anyhow, these represent only two of the several definitions available.

As a matter of fact, many stakeholders, among which investors, banks, financial service providers, regulators, policy makers and academia are trying to explain what 'green' means, resulting in definitions that vary in terms of scope, level of detail, transparency and other dimensions.

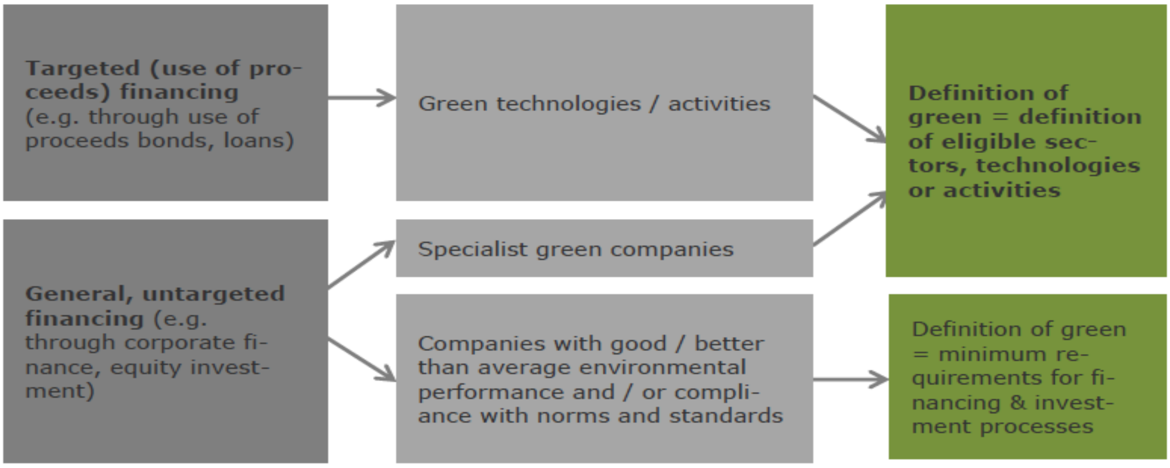
Moreover, 'green finance' is typically defined taking into consideration what it finances, meaning investments into green technologies, activities and companies, and not what it gets, like investments which bring an environmental impact. However, focusing only on what is financed, without considering how it is financed, does not take in consideration the mechanisms for which investments products may affect the environmental impact of companies in which they are invested.

Indeed, it is important in order to make everything clearer, to make a distinction between targeted and untargeted finance/financing, like in the scheme n° 1:

---

[2] http://www.pbc.gov.cn/english/130721/3133045/index.html

Scheme n° 1[3]



- Targeted financing: use of capital for growth and application of green technologies/projects or companies whose revenues are mostly generated thanks to green activities/technologies. Examples are green loans, green bonds or green project finance;
- Untargeted financing: use of capital for companies that manage very well environmental, social and governance (ESG) risks, making them more environmentally friendly than others. Examples can be found in private equity and listed investments or in corporate finance. In this way, it is possible to reduce the environmental risk in investors' portfolios and increase the awareness of the importance of ESG risk management, moving companies towards green development.

[3] Kahleborn, Walter, A. Cochu, I.Georgiev, F. Eisinger, D. Hogg, *Defining "green" in the context of green finance*, Luxembourg Pubblication Office of the European Union 2017, pag 3

1.1.1 The Need of Clarity in Green Finance

The 2016 G20 Green Finance Synthesis Report states that "In many countries and markets, the lack of clarity as to what constitutes green finance activities and products (such as green loans and green bonds) can be an obstacle for investors, companies and banks seeking to identify opportunities for green investing."[4]

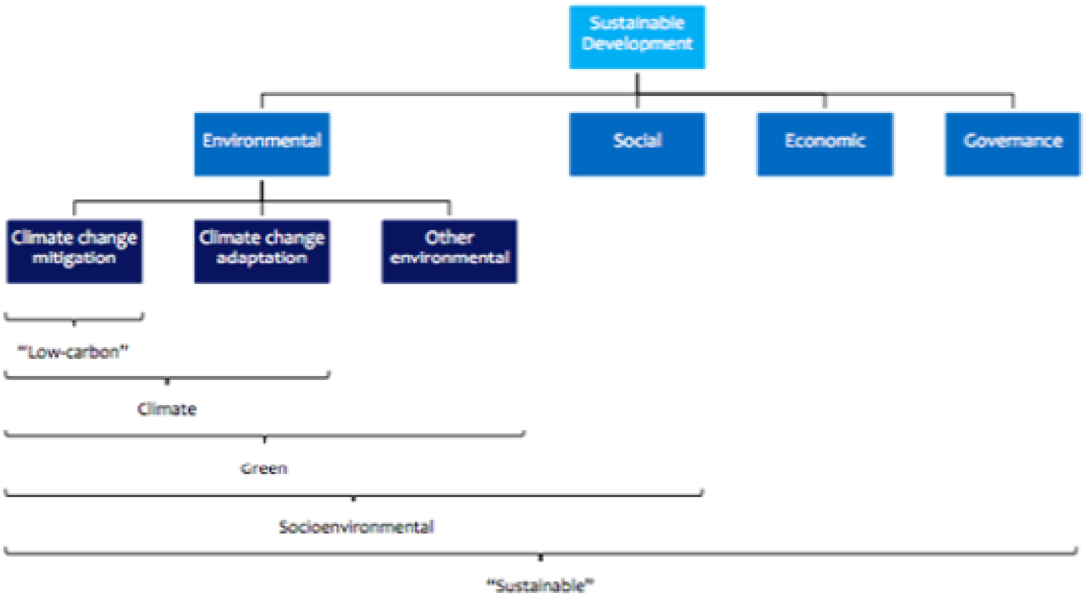
In effect, it happens that the terms green, climate and sustainable finance are used interchangeably, but they are nested concepts.

Sustainable finance deals with the integration of environmental, social and governance factors across financial system, with the purpose of strengthening resilience, targeting capital allocation and improving accountability.

In the scheme n° 2 the classification of green, meaning environmental, activities is proposed.

[4] The need of common language in Green Finance. Towards a standard-neutral taxonomy for the environmental use of proceeds, pag 11.

Scheme n° 2[5]



The same could be done also to social, economic and governance activities.

Indeed, the need of a 'universal common taxonomy', a set of minimum standards and disclosure frameworks is evident, in order to classify green activities underlying financial products and to efficiently allocate financial resources. It represents a priority for many stakeholders, not only to focalise finance on particular green products or players, but also to highlight changes needed on an economy-wide level. Moreover, it will be possible to make descriptions and comparisons and to develop standards and labels to direct the choices of market participants.

Furthermore, this will enable to avoid confusion within investor community, preventing from inconsistencies and overlap with already existing requirements.

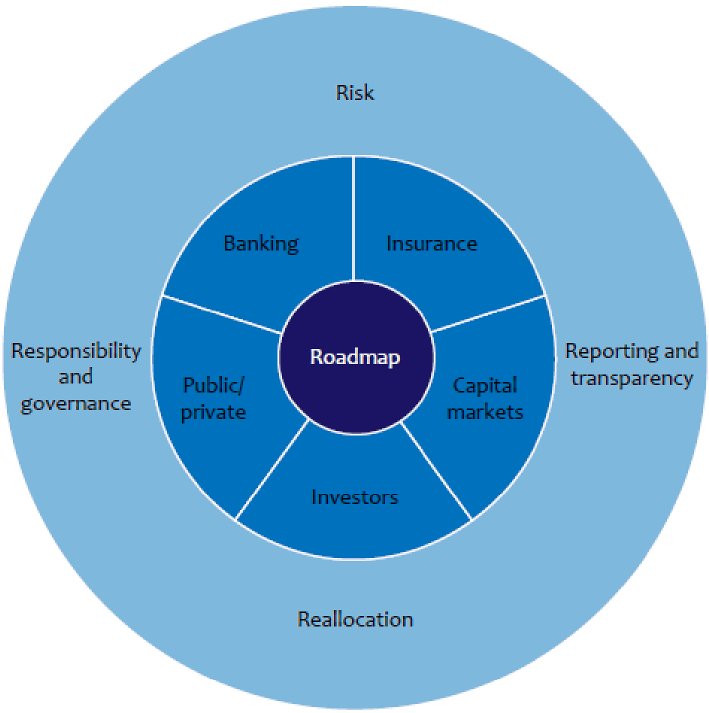
[5] *UNEP Inquiry* 2016

Finally, a harmonization process will avoid the phenomenon of greenwashing, consisting in over-marketing the environmental factors, since lack of sanctions, controls and knowledge in the proper assessment of investment policies.

1.2 The 5 R's for A Sustainable Financial System

It is important to pay attention on five common themes, which pass through the fundamental sectors of finance, that are essential for the establishment of sustainable financial systems. They are reported in the scheme n° 3[6]. Specification and explanation are required.

Scheme n° 3



[6] *Financing the Future. Report of the Italian National Dialogue on Sustainable Finance,* pag 14

The first element to analyse is reallocation of capital for mainly three fields:

- Green finance: it may result often difficult for small and medium-sized companies to access green finance. Financial innovation such as crowdfunding and peer-to-peer lending may represent a solution;
- Sustainable infrastructure: both private and public finance for infrastructure are not sufficient and not connected to sustainability priorities. OECD has provided policy recommendations in the way of channelling institutional capital for green infrastructure;
- Clean tech innovation: bettering quality and quantity of innovative finance is essential for themes like smart buildings and cities, clean energy, improved food yields and less food waste, universal access to education and health and monitoring of critical ecosystem.

The second one, instead, is risk since sustainability factors represent more and more considerable factors of risk for financial institutions. The awareness that traditional systems are not enough to deal with environmental threats is growing.

Three types of risks have been recognised by the Bank of England:

- Physical risk: direct threats from natural and man-mad hazards, such as pollution or storm damage to buildings;
- Liability risk: indirect threats arising from fights with financial institutions for activities that might harm the environmental sphere;

- Transition risk: indirect threats resulting from responses to environmental degradation, such as changes in policies, technologies and consumer preferences.

The third component is responsibility, which is becoming part of fiduciary duty and institutional investors' obligations. It represents also a driver for market innovation, and it will lead financial intermediaries to better their capabilities and skills through trainings.

Reporting is the fourth theme to deal with and it must be done not only by corporations, but also by financial intermediaries to allow consumers to understand if ESG risks are properly taken into considerations. The new FSB Task Force on Climate-related Financial Disclosure, sponsored by the Financial Stability Board, gives guidelines for this purpose, stressing the importance of a forward-looking disclosure beyond past performance one. Moreover, in this way investors will make informed choices, clients will choose the proper financial product and regulators will estimate threats to sustainable financial system.

Lastly, roadmaps are needed, since most countries own sustainable financial system elements but they are not connected strategically.

1.3 The Current International Situation

In order to understand the impact and the growing importance of green labelled products, a general overview is need on the actual global situation.

The Global Sustainable Investment Alliance (GSIA) deals with providing, every two years from 2014, the so called Global Sustainable Investment Review, consisting into a report which collects results from the market study of regional sustainable investment forums from Europe, United States, Japan, Canada, Australia and New Zeeland. It provides information regarding the SRI assets, meaning social responsibility investment assets, even if due to the lack of a clear definition of green finance, investment flows and associated impact might not be assessed properly every time.

Last data available date back to beginning 2018, with an amount of global sustainable investment of about $30.7 trillion in five major countries, with an increase of 1.34% from 2016. Details are reported in the tables n° 1 and N° 2:

Table n° 1[7]

Region	2016 ($ bln)	2018 ($ bln)
Europe	12.04	14.075
United States	8.723	11.995
Japan	0.474	2.18
Canada	1.086	1.699
Australia/New Zeland	0.516	0.734
Total	22.89	30.683

[7] *2018 Global Sustainable Investment Review*, Global Sustainable Investment Alliance, pag 8

Table n° 2[8]

Region	2014	2016	2018	Growth Per Period	
				Growth 2014-2018	Growth 2016-2018
Europe	€ 9,885	€ 11,045	€ 12,306	12%	11%
United States	$ 6,572	$ 8,723	$ 11,995	33%	38%
Canada	$ 1,011	$ 1,505	$ 2,132	49%	42%
Australia/New Zeland	$ 203	$ 707	$ 1,033	248%	46%
Japan	¥ 840	¥ 57,056	¥231,952	6692%	307%

The first one collects the amounts of global sustainable investing assets expressed in billions of US dollar, converted at the exchange rates at the time of reporting. The second one, instead, represents the amounts in their own currencies with the percentage of growth from 2012 until 2018. What has to be highlight is that the amount of assets is continuing to grow, more in some regions than in others. While United States' increase is higher than previous year, elsewhere the rate of growth is lower than previous biennial.

Except for Europe, the proportion of sustainable assets relative to the total managed assets grew in all regions, as can be visible in the table n° 3.

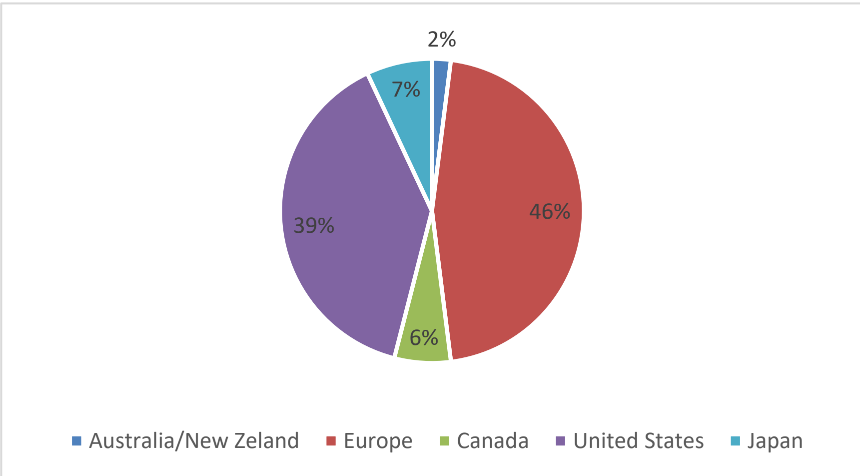
[8] *2018 Global Sustainable Investment Review*, Global Sustainable Investment Alliance, pag 8

Table n° 3[9]

Region	2014	2016	2018
Europe	58.8%	52.6%	48.8%
United States	17.9%	21.6%	26.7%
Canada	31.3%	37.8%	50.6%
Australia/New Zeland	11.6%	50.6%	63.2%
Japan		3.4%	18.3%

At least, dwelling with who is in the position of having the majority of sustainable investing assets at the beginning of 2018, Europe won the title, as outlined in the graph n° 1.

Graph n° 1[10]



1.3.1 The Situation in Italy

Due to the lack of a unique taxonomy to define 'green finance', it is difficult to have a proper update on the current state. Nevertheless, the situation in Italy in 2016 was

[9] 2018 *Global Sustainable Investment Review,* Global Sustainable Investment Alliance, pag 9

[10] Ibidem

the following: 27.5% of companies could be defined ‘core-green’, i.e. producing high environmental valued goods and services, 14.5% ‘go-green’, i.e. having an environmental management system and using high quality environmental standard in the process of production.

Its industry performance index, based on 16 indicators, was 59.4 on a scale from 0 to 100, where 0 and 100 represent respectively the worst and the best performance for all indicators.

The share of retail sustainable funds offered by Italian asset managers reached a record level in 2017, despite only few important institutional investors lead the market. Retail investors, for a percentage of 45, showed their interest to invest in SRI. However, financial education is needed for retail investors and financial advisors in order to improve their offer and advisory services on SRI financial education.

Moreover, Italian pension funds have demonstrated an increasing commitment to SRI, even if it is not enough to decrease the gap with other Europe SRI markets.

1.4 The Steps Toward an Environmental System

Over the past decade, there have been increasing efforts to align financial system to long-term sustainable development and there is a growing recognition that ESG factors are fundamental for creation of value.

Environmental threats are creating risk to financial assets and new opportunities for insurance companies, while banks, institutional investors and capital markets are incorporating environmental and social factors in decisions for capital allocation.

Through the fundamental position of the public finance in allowing this shift, the private sector should bring most of the capital requested.

Three agreements, reached in 2015 by the governments of the world, have been fundamental for the coming decades:

- Sustainable Development Goals: the centrepiece of the 2030 Agenda for Sustainable Development, approved by 150 international leaders met at the United Nations to contribute to global development, promote human well-being and protect the environment. They amount to 17 goals and 169 sub-objectives, which aim to shift capital away from damaging ‘business as usual’ trends towards end to poverty, fight against inequality and in favour of economic and social development;
- Paris Agreement on climate change: it brings together all nations to combat climate change to keep the global temperature rise below 2° C above pre-industrial levels, with the aspiration to limit the temperature increase to 1.5° C, and in addition, to improve the ability of countries to react to climate changes. In order to do these, new technology frameworks and enhanced capacity building frameworks are needed, supporting countries’ actions in line with their own national objectives;
- Financing for Development: outcome document resulting from Addis Ababa Action Agenda conference, which focuses on the steps to enhance domestic and international resource mobilisation for developing countries in terms of public and private capital.

1.4.1 Improvements Needed

In order to reach all the targets established at the speed and scale required some elements need to be met:

- Policy action: to eliminate market failures, like unpriced pollution and resources. Even if improvements have been made in terms of internalisation of externalities into market prices, market failures in effectively pricing scarce resources remain, resulting in adjusted returns for sustainable finance inadequate to attract enough capital;
- Public finance: essential both domestically and internationally to provide goods that the market cannot deliver, to stimulate private actions thanks to subsidies and incentives and to assist the transition process of developing countries. It is fundamental in order to identify cost-effective options with leverage impacts;
- Market and institutional barriers: actions are necessary to eliminate barriers which can represent an obstacle to the proper allocation of capital to sustainable development. Examples can be short-terrorism, inadequate risk management, misaligned incentives, lack of transparency and poor stewardship. Actions are important for critical financial challenges, among which:
  - Reducing capital: the use of natural capital with clean technologies and human expertise for a sustainable development often requires higher

upfront costs, but at the same time lower operating ones, resulting in a better life cycle costing;

- Speed and scale: levels of sustainable finance must be increased as fast as it is possible in order to achieve time-bound targets;
- Extending time horizon: for achieving short-term goals over the next decade, a strategic view is needed, especially for long-lived buildings.

1.5 The Role of Banks

Bank credit amounts around two thirds of investments, representing the principal financing tool in the European Union economy. Even if the banking sector plays the first role in financing investments, only a small part is explicitly considered ‘green’.

There is a strong debate as far as the role of banks in promoting and facilitating the shift to a more environmental financing. The majority confirms the role of the banking sector as the main operator toward an environment system thanks to its essential role in helping the capital flows through lending, investment, advisory and supporting innovative technologies. Others, instead, believe in a limited role, due to the different time horizons, i.e. the mismatch between the short-medium nature of the banking sector for investment and risk performance management and the longer one of sustainable development actives and so environmental risk, creating other financial stability problems.

There is the need, and in this consists the challenge of banks, to bridge supply and demand, considering the whole spectrum of risks, integrating also environmental and climate change (ECC) risks.

Furthermore, banks impact with their activities also assisting costumers in transiting to more environmental business model, beyond considering environmental problems in its own operations and in the financing of portfolio.

An effective public-private cooperation is peremptory in order to speed up the achievement of a sustainable development and more needed to be done in order to attract private capital.

Moreover, since the uncertainty of regulatory environment forces and the need to be compliant with regulatory requirements, banks are forced to a constrained version of balance sheets, that must comply with the current regulatory framework, optimise capital according to actual regulatory environment and prevent to lock-in capital for the possibility of changes in future regulatory framework.

1.5.1 Examples of Banking Support

Banking activities represent a valid support to promote and finance environmental system both for public and private sector, with strong goals in terms of reduction of carbon-related industries and increase of renewable energy ones.

A wide range of environmentally friendly products is provided by banks, focusing the attention on both the quality and the innovation that better sustainability.

Examples sectors in which they operate are provided below:

- In the retail sector, they provide loans for environmental purpose, like installation of photovoltaic or solar panels, the purchase of ecological equipment and green mortgages, in terms of acquisition of energy efficient properties or loans for their energy-efficient renovation;

- In the corporate sector, they provide business loans for photovoltaic panels hydropower or biomass plants. In the agricultural field, services are provided for covering damages to the crop in case of atmospheric events like rain and hail, or in case of a fire;
- Private equity firms within banks have introduced ESG criteria into their investment decisions or use ESG key performance indicators, i.e. KPIs, in order to monitor clients' portfolios.

2. Green Mortgages and Real Estate Market

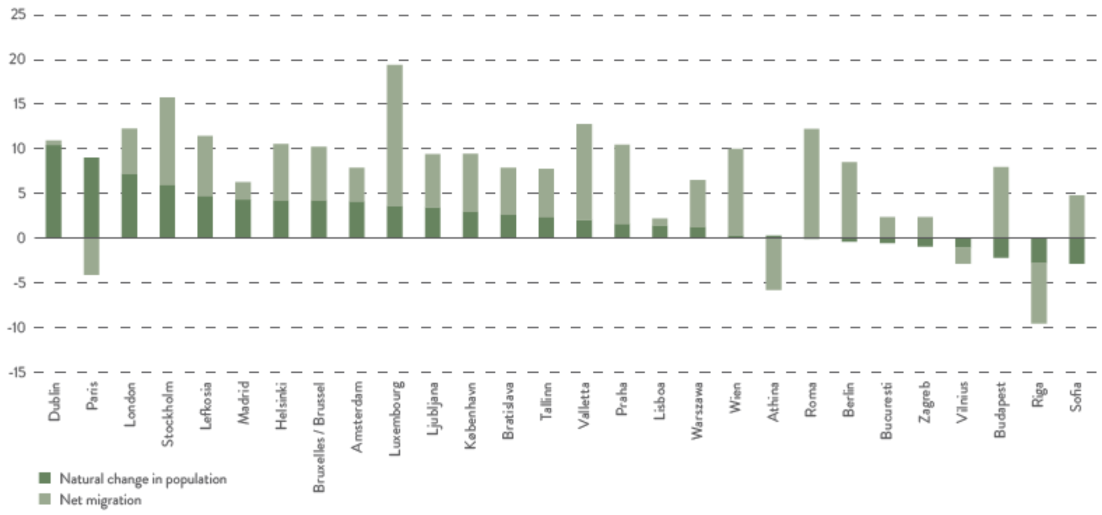
This work, from now on, will focus on only one type among green financial instruments: green mortgages. Nevertheless, before providing a definition and explanation on the relevance that they have nowadays, it is appropriate to give a general outlook of firstly the European real estate market and subsequently the Italian one, since it is the one in in which the analysis will be performed.

2.1 European Housing Market

One of the phenomena which homeowners have to face in Europe is the so called 'urban paradox', i.e. the choice to live in city, with the consequence of having lower space and higher costs, rather than in more rural area offering more space at lower costs. Moreover, even if cities bring more opportunities, they are characterised by problems like social inequalities, crime, higher unemployment and so on.

Nevertheless, there is a general increase in terms of population in mostly of capital cities in European Union, with an average of 7 people per 1000 inhabitants in the last ten years, which is highlighted in the graph n° 2:

Graph n° 2[11]



Of course, this phenomenon varies according to the distribution and the morphology of cities in Europe, which depends on economic, historical political and administrative reasons, and the importance of capital cities in the respective countries. Italy and Germany, for example, are based on a polycentric pattern of urban development with many cities and Rome and Berlin, respectively, contribute less than 10% to GDP. United Kingdom and France, instead have a more monocentric urban development, with London and Paris accounting for 30% of GDP.

However, independently where the trend is large or small, 3 out 4 people in EU live in urban areas, which attract mostly the younger population and migrants.

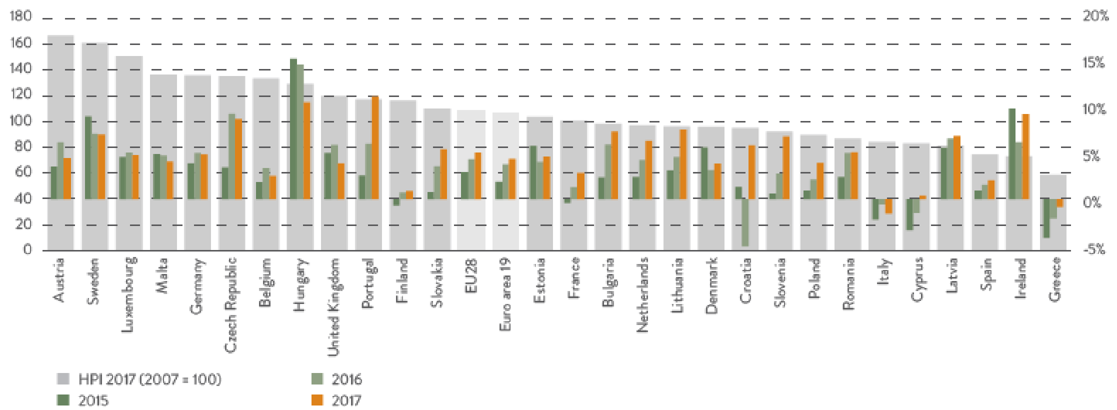
As far as house prices, according to last data available of 2017, half of European Union countries have reached or gone beyond the threshold of the crisis year, while

[11] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 28

the remaining part have remained below it, even if the majority of those countries have reached 80-90% of the level registered in 2007.

Anyway, there is a general ascending trend of house pricing, 6% with respect previous year, with a strong acceleration in 65% of countries, which is visible in the graph n° 3.

Graph n° 3[12]



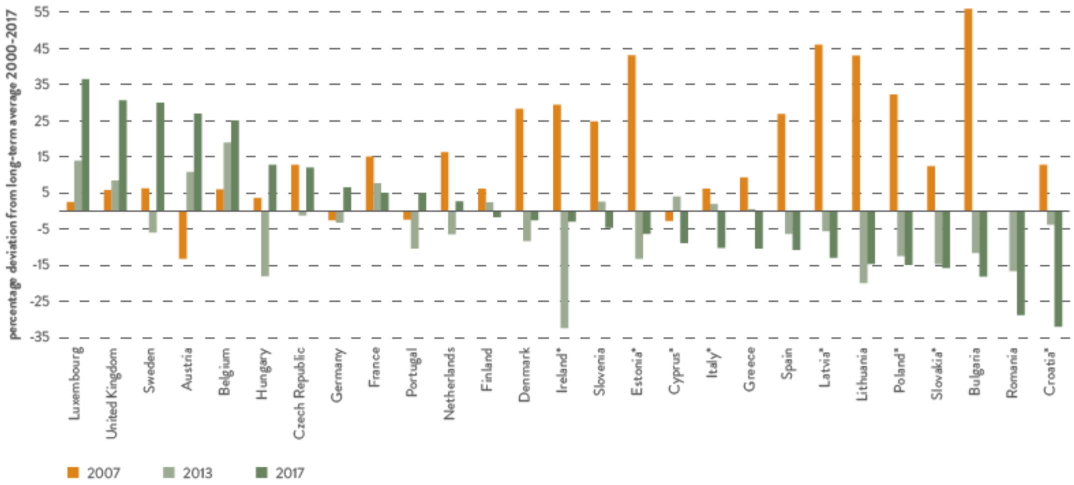
This behaviour in prices is obviously reflected in non-affordable homes, as demonstrated by the housing overburden rate, i.e. proportion of residents whose housing costs go over 40% of their disposable income net of housing allowances. In these costs there are included, beyond rent or mortgage payments, utility bills, maintenance costs, insurance and taxes.

Moreover, comparing the house price index and the per capita gross disposable income through a ratio, i.e. the price to income ratio, a measure of house affordability is obtained. Since it is not possible to perform a cross-countries analysis, because

[12] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 31

it will result statistically not significant, the deviation from the long-term average of the price to income ratio is used. Data are taken from 2000 to 2017, since for previous period some can be missed.

Graph n° 4[13]



In the graph n° 4, countries are ordered according to their deviation from the respective long-term mean in 2017. It is possible to note that, on average, there is a tendency of an improving in the ratio, with respect 2013, corresponding to a lower affordability capacity.

Beyond the enhance of unaffordability, the increase in prices, together with the one in mortgages, lead to an improve in construction investments. As a matter of fact, the construction investments enhanced by 3.5% in 2017, continuing their growth path, even if pre-crisis level is not yet reached. The same considerations can be for

13 *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 32

issuance of building permits, which are a proxy for construction investments, and for residential transactions involving both new and existing houses.

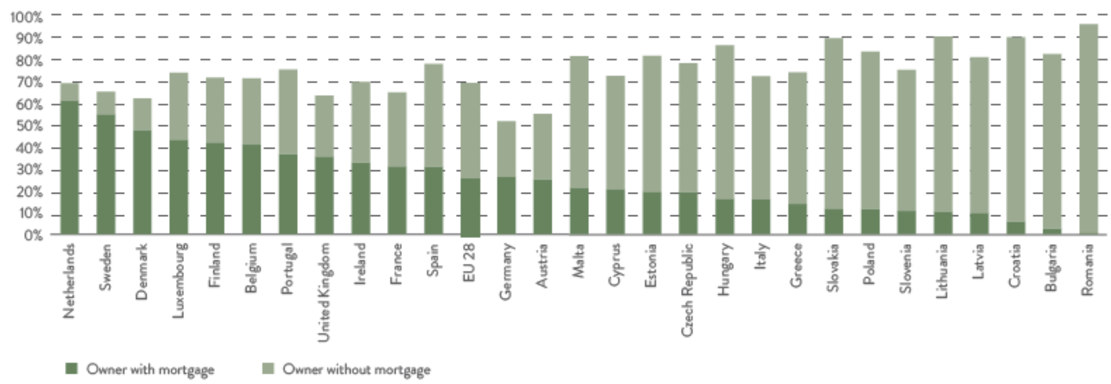
Although a positive boost of supply side has been registered, it is not enough to satisfy the lack of houses, mainly in area where the demand is the highest. Moreover, the existing housing stock is unsuitable for consumer needs, resulting in overcrowded or under-used dwelling.

2.2 European Mortgage Market

After having analysed the situation regarding the housing market, it is worth to put the attention on the mortgage one.

The concentration of mortgages is heterogeneous in European Union and this depends not only on the economic factors, but also on cultural aspects, like the preference of people to invest in homes or in other assets. This is pointed out in the graph n° 5.

Graph n° 5[14]



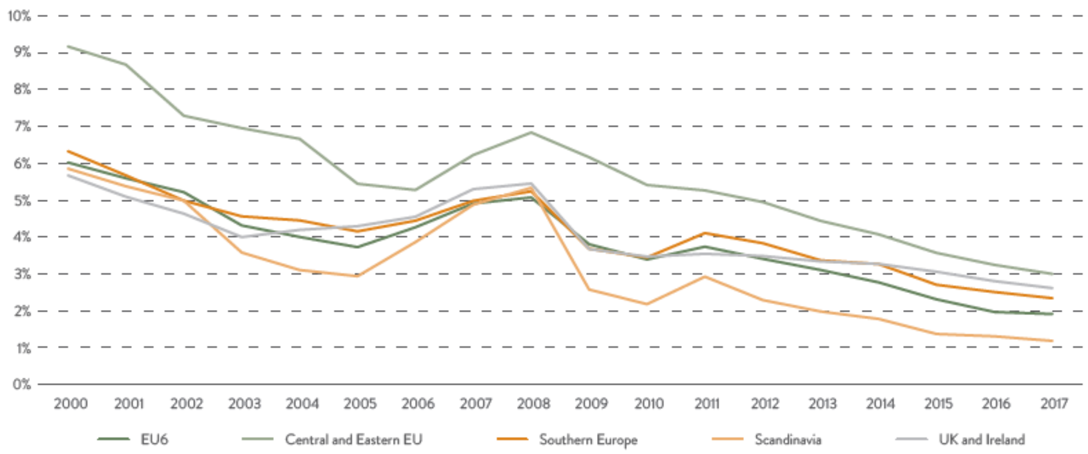
[14] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 35

Anyhow, independently by owning or renting a house, the mortgage market represents an important piece of the European wealth, reaching 46% of GDP in 2017, with an enhance in aggregate term of 2.3%. Nevertheless, UK, Germany, France, Netherland and Spain accounts for the majority of outstanding mortgages, more precisely 72%.

The average mortgage amount for an adult has increased by 3.2% than 2016, reaching 16,872 €, 2,000 more than in 2007, showing how it follows the positive path of prices. It seems that since dwellings are more expensive, borrowers are borrowing more.

As far as the mortgage interest rate, it has proceeded its decreasing trend, lowering the previous year amount of 22 bps and reaching 2.42%, with details on the different areas in the graph n° 6.

Graph n° 6[15]



[15] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 39

The descending path of interest rates is accompanied by greater preference in longer initial fixed mortgage rate period to lock-in the interest rate, to avoid possible future increase.

2.3 The Italian Situation

The Italian outline, instead, is not always in line with the average outlook of the European Union.

2.3.1 The Housing Market

Italy experienced a decline in terms of house pricing in 2017, precisely 0.4% with respect previous years, bettering the sales performance. This is due by the compensation in the raise in the price of new dwellings of 0.1% with the decline in the price of existing ones of 0.6%.

Houses transactions, instead, experience the same growth of 2016, 4.9%, for an amount equal to 543,000; in the table n° 4 details for macro area.

Table n° 4

Macro Geographical Areas	Dwelling Transactions' Increase %
North- East	5.3
North-West	4.9
Centre	3.5
South	5.8
Islands	4.9

Data available from the last ISTAT census have showed that residential buildings amount for 12.2 million with over 31 million dwellings, with 74% corresponding to multi-family buildings and 24% to single-family ones.

Regarding the oldness of homes, 60% have been built before law n° 373 of 1976, i.e. the first law at national level for energy saving. The age doesn't not exactly entail a poor state of preservation of building fabrics, even if 2 million properties were considered by ISTAT in bad or poor state of conservation, respectively 1.75% and 15%.

Moreover, the average amount for electricity energy and heating needs for residential buildings ranges from 160 to 180 kWh/m^2 per year, while in case of houses which are over 45 years old the amount shifts from 160 to over 220 kWh/ m^2 per year.

2.3.2 The Mortgage Market

As the great growing path in 2016 (27.3% more than 2015), the issuance of new loans for residential houses decreased in 2017, with an outstanding amount of € 375,4 bn, but the reduction in prices and enhancement in the Italian consumer confidence index are valid elements in favor of a growth in the future.

However, 50 % of total of houses purchased are linked to a mortgage, with an average amount of € 126,000, with an average maturity of 26 year and with the prevalence of fixed-rate; the last is motivated by a preference in a certain constant rate due to the fear of a possible increase since interest rates at historic lows. Indeed, the interest rate on short term fell to 1.53%, for a maturity over 1 year stayed constant to 2.12%, while the one for new transactions declined to 1.90% (in 2007 was 5.72%).

According to the 'Italian Statistical Yearbook', the average monthly installment for mortgages is € 586.41 while the one for rent is € 430.5.

Also, not-performing loans decreased from € 65.9 bn to € 64.4, with a percentage of 25.8.

2.4 What Are Green Mortgages?

The term green mortgage is used to refer to loans for financing residential, commercial and retail real estate, both for purchase of new energy efficient dwellings and for the renovation of existing ones investing in energy efficient components. They include preferential interest rates, rebates and tax advantages.

People are facing increasing costs of energies and low energy efficiency in their properties, with the subsequent lower monthly disposable income to pay debt obligations. Besides, the loss of energy subsidies will concern low- and middle-income class households as energy bills represent higher percentage of monthly cash and they live generally in lower quality houses. Moreover, poor quality building increases expenses for reparations and the risk of devaluated dwellings if have to be repossessed by the bank, resulting in higher cost faced by homeowners for borrowings. Furthermore, bad construction impacts negatively the environment in an increasingly carbon-conscious society. Finally, the introduction of higher standards for energy efficiency due to environmental and economic drivers, threats banks and borrowers, since 97% of building in Europe are considered inefficient and the rate of energy renovation needs to pass from 1% to 3% if Europe wants to reach its climate targets.

Indeed, green mortgages may represent a possible solution, since they can be seen as an opportunity for both lenders and borrowers. On one side, lenders take advantage of a lower probability of default, enhancing in this way the value of homes and offering a lower cost of financing, meaning lower interest rate, but keeping profit margins. On the other hand, borrowers have greater purchasing power to be exploited for better quality property since minor energy, repair and health costs can be recorded, decreasing the monthly expenses relative to standard ones. In this way, investors and developers could transform the real estate market toward a low carbon economy.

This mutual benefit represents a possibility to positively modify the standard of millions of properties. This will be plausible only if green mortgages become a mainstream offering available for every borrower. They could give a substantial boost to climate, energy and risk goals only if there will be a strong collaboration among various participants. This means, stronger demand by consumers, business case for banks to provide these instruments and regulators offering incentives to sustain the market.

To notice that the most cost-effective period for financial institution to assist in energy bettering of buildings is when properties are built or bought, in the earliest phase of transactions.

2.4.1 Characteristics of Green Homes

With the term green homes people refer to houses designed and built with a focus on minimizing or eliminating, if it is possible, the impact on the environment. Certifications are required in order to consider a home green, and these differ across countries.

Among characteristics that need to be taken in considerations, some are provided below:

- Construction site: referring to steps to avoid damaging the surrounding environment;
- Location: in terms of proximity to public transportation (railway or bus stations) or in a walk-able area, those reducing transportation impact;
- Energy efficiency: in terms of better insulation, depending on doors and windows, heating and better air conditioning systems;
- Sustainable and healthy materials: meaning non-toxic materials for occupants and for manufacturers, with heavy ones coming from closed factories to lower transportation costs. People need to opt for durable materials, since reduce the frequency and severity of repairs, and prefer recycled, up-cycle or re-purpose items;
- Lighting and shading: meaning orientation of property in the best way to collect as many as energy possible in winter and to protect from the sun in summer;

- Indoor air quality: in terms of natural ventilation or technologies resolutions for healthier air.

Among categories that may prefer green mortgages rather than standard ones, younger generation and middle-class borrowers stand out.

Nevertheless, only few borrowers have showed strongly interested in this type of financial service. Many reasons can be reported, but the most certain is the lack of awareness around green finance.

Among financial institutions who have already offered these products, reasons in doing so can be summarised below:

- Commercial: proposing a competitive instrument to satisfy the needs of clients to get stronger relationships with them and to diversify investors' portfolios;
- Regulation: government support for bettering credit and asset quality;
- Corporate social responsibility: promote energy savings and support and strengthen social responsibility.

Nevertheless, many intermediaries have not been able to provide these services, since barriers for access to green market are present:

- Interest: low interest among borrowers;
- Operational: lack of green filtering in IT system and not favourable legal framework. As result, financial institutions are not able to create green cover pool, allowing them to issue 100% of green bonds, but they have to mix their

cover pool with green mortgages, not enabling them to take advantage of lower funding costs;

- Technical: lack of expertise to transfer energy efficiency into green value.

Changes have to be made both on the demand and supply side. As far as the first one, it will depend from homeowners and commercial property owners, who need clear and compelling benefits from these mortgages with respect standard ones and that benefits of course should outweigh expenses. Concerning the supply side, instead, the attractiveness for lenders, coming from reduction in financial exposure, need to be proven and long term, through lower default rates, better loan to value ratio and lower losses for them in case of borrowers' default.

Finally, investors regulators and policy makers may help in growing the interest in this topic. In effect, with investors asking for more green assets, lenders will have a ready source of funding for energy efficient mortgage, instead of using capital markets as usual. Providing a standard definition of energy efficient mortgages with clearly defined building performance assessment will represent a great support towards this path. Nevertheless, to do so, regulators need to be sure that these products are useful to reach climate change goals.

2.5 Energy Efficient Mortgage Initiative

The European Union has imposed itself targets in terms of energy savings for both 2020 and 2030, respectively 20% and 30%. Taking advantage of mortgage financing benefits to support energy efficiency can be a great help for these goals, since

European Mortgage accounts for 53% of EU's GDP and properties are responsible for 40% of energy consumption and 36% for $C0_2$ emission.

Towards this idea, the European Mortgage Federation-European Covered Bond Council, i.e. EMF-ECBC, has been working on a project known as 'Energy Efficient Mortgage Initiative', which for the first time pull together major banks in quality of mortgages lenders, data providers, companies coming from construction and energy field to debate on private financing including energy efficiency.

Two are the main reasons for which the project has started: on one hand the awareness of the tactical role of financial institutions to assist in reaching targets above mentioned by bring the problem to consumers and providing standardised approach for financing energy efficient buildings or energy efficient renovation; on the other hand, a response on the increasing demand of green funding side, by having this new asset class available to issue green bonds and green covered ones.

However there is one fundamental hypothesis behind the project as already stated in this work: more energy efficiency will bring to lower utility costs and as consequence lower probability of default of borrowers, improving their probability to repay the loan, since more disposable income will be available, and at the same time enhancing the value of properties with the result of decreasing the loss of banks if default will happen.

Moreover, by applying these risk mitigation factors in the calculation of capital requirements for banks and a positive confirmation is stated, a realignment in the requirements of capital will be recognised by the regulatory framework.

The Energy Efficient Mortgages Initiative is formed by two projects: the ‘Energy Efficient Mortgages Action Plan’ (EeMAP) and the ‘Energy Efficient data Portal & Protocol’ (EeDaPP).

The first project has the following purposes:≪ (1) identification and summary of market best practices, (2) definition of an energy performance indicators and a Building Energy Passport, (3) identification of pre-requisites for the assessment of ‘green value’, (4) substation of correlation between EE & probability of default-portfolio analysis and (5) definition and design of energy efficient mortgage, based on preferential financial conditions≫[16].

The second one, instead, brings the following aims: ≪ (1) Identification and energy summary of market best practices within data systems, (2) definition of energy efficiency reporting criteria, (3) design and delivery of standardised data protocol and common centralised portal, (4) data and substitution correlation analysis, and (5) roadmap for system integration≫[17].

[16] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 20
[17] *Hipostat 2018 A Review of Europe's Mortgage and Housing Markets*, European Mortgage Federation, pag 20

3. Statistical Methodologies

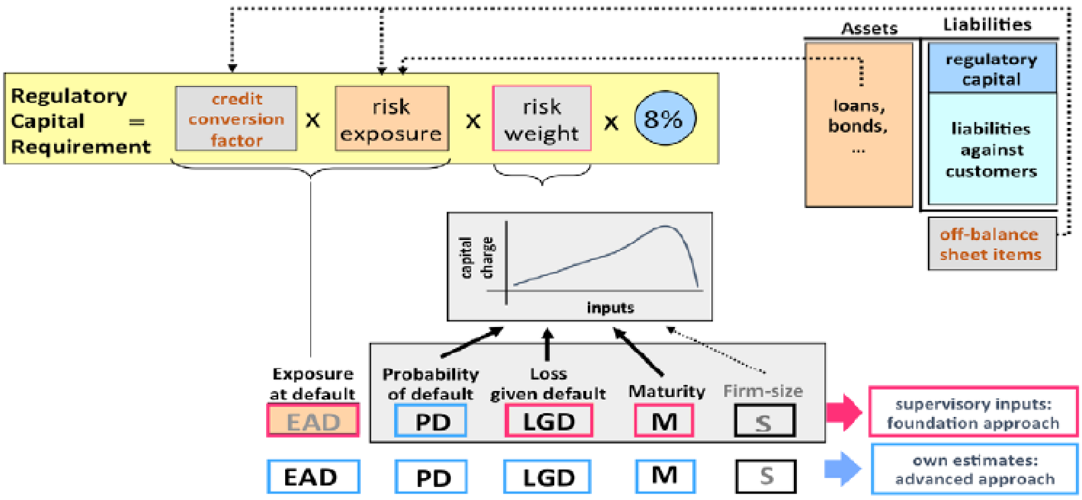
While previous chapters where focused in giving, firstly, an idea on the concept of green economy and, secondly, a general background on the European residential and mortgage situation with a deeper concern on the Italian one, from now on the analysis will convert to the empirical side.

The purpose of this work is to assess if there is a relationship between the energy performance of a building and the corresponding credit risk.

3.1 The Bank's Approach to Risk

Banks have to accurately estimate all the potential risks related to their activities and allocate a minimum prudential amount of capital, to face possible losses, to hold in their balance sheets. This quantity can be assessed by a risk weighting process, involving risk-weighted assets, as described in scheme n° 4.

Scheme n° 4[18]



All different risks, consisting in credit risk, counterparty credit risk, market risk and operational risk, need to be valued in accounting terms and measured using either a Standardised Approach (SA) or Internal Ratings-Based Approach (IRB).

Under the SA, there is no way for now to consider also energy performance in the calculation of capital requirements, due to the rigid and abstract manner to calculate credit risk. Under IRB approach, instead, other measures are used, like probability of default (PD), loss given default (LGD), maturity or size, which can be taken in consideration to elaborate a correlation with the energy performance characteristics of the asset to be weighted.

This work will focus on the probability of default as dependent variable, among just mentioned ones.

[18] M.Billio, M. Riedel, A. Bedin, D. Leboullenger, *Technical Report on the Methodology Design to Carry Out Portfolio Analysis,* L. Bertalot, pag. 7

3.2 Correlation or Causality Analysis?

Correlation and causality are not the same thing; a moment need to be taken to linger on the difference between the two.

The idea behind, as it was previously mentioned in this work, is to demonstrate that green properties are less risky than not efficient ones, since the borrowers' probability of default will decrease due to more disposable income available in order to repair mortgages, thanks to lower utility expenses to face, resulting in advantages for both financial institutions and investors (households and companies).

Correlation allows to analyse the relationship between two events; it can be positive or negative, meaning that when an event occurs the probability of the other to occur is high or low respectively.

Causality, instead, deals with identify origins and causes to demonstrate that one event directly affects a change in the other one.

Unfortunately, a high level of historical data is required to perform a causality testing, but this is not feasible due to the contemporaneity of energy efficiency and performance data in the financial system.

Indeed, only a correlation analysis can and will be carried out in this work, even if the use of longitudinal data which can catch both time and cross-sectional variation represents a valid alternative to provide a comprehensive one.

Nevertheless, the causality study will be a positive enhancement in the future for an academic point of view when a data set with a significant depth will be available.

3.3 Statistical Methodologies Applied

This section will treat the methodologies used to perform the credit risk study: Logistic model firstly and Cox Hazard model secondly. Both are multivariable models which express the relation between two or more predictors (independent) variables and one outcome (dependent) variable, where the predicted value of the outcome variable is seen as the sum of products, each one obtained by multiplying the value of independent variables with their respective coefficients, which indicate the impact of the independent variable on the dependent one adjusting for all the other independent variables.

While in the logistic regression the variable is usually a binary event, in the hazard model it is the duration of time to the occurrence of a binary failure event during a follow-up period of observations.

3.3.1 The Logistic Regression

The logistic regression is a mathematical modelling approach used to explain the relationship between several continuous or categorical independent variables and a dichotomous dependent one.

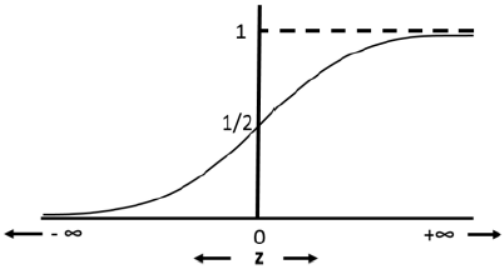
Due to its simplicity, it is one of the most popular and used models. In particular two are the predominant reasons: (1) the estimates ranging from 0 to 1 of the logistic function from which it derives and (2) its S-shape, as it is immediately explained.

The logistic function $f(z)$ is defined by the formula

$$f(z) = \frac{e^z}{1+e^z} = \frac{1}{1+e^z}$$

and takes value ranges from 0 to 1, regardless the value of z, as it is possible to notice on the graph n° 7.

Graph n° 7[19]



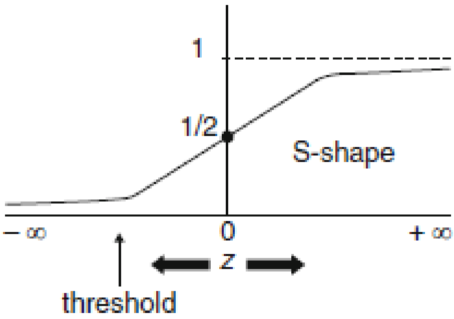
As z approaches $-\infty$, $f(z)$ goes to 0, while as z approaches $+\infty$ $f(z)$ goes to 1.

Indeed, since logistic model is designed to estimate a risk that stands between 0 a 1 and probability always ranges between the two values, and other models do not always provide this result, it is the first choice when probability need to be estimated.

As far as the second reason, the logistic function follows a S-shaped curve. Actually, it will start with slow and linear growth, followed by an exponential increase which then slows again to a stable rate as it is visible on graph n° 8.

[19] M.Billio, M. Riedel, A. Bedin, D. Leboullenger, *Technical Report on the Methodology Design to Carry Out Portfolio Analysis,* L. Bertalot, pag. 21

Graph n° 8[20]



The risk firstly is minimal at low values of z until some threshold, then it increases rapidly for a large range of z values and then remains extremely high around 1 once z gets large enough. In this way, it is possible to combine the effect of several risk factors on the risk for an event.

Now that the logistic function has been explained, the next step is to see how arrive to the logistic regression.

It is enough to express z as the sum of a linear combination of p covariates x, i.e.

$$z = \alpha + \beta_1 X_1 + \beta_2 X_{21} + \cdots + \beta_k X_k$$

and substitute it in the exponent of the logistic function, obtaining

$$P(Y^i = 1 \mid x_1, x_2, \dots, x_p) = \frac{1}{1 + e^{-(\alpha + \sum_{l=1}^{p} \beta_l x_l)}} = G(\beta^i X^i),$$

where α and βs are coefficients to be estimated. In this way, the function $f(z)$ can be expressed as the conditional probability of a binary dependent variable Y of a subject i given observed independent variables $x_1, x_2, \dots, x_p$.

[20] D.G.Kleinbaum, M.Klein, *Logistic Regression, A Self-Learning Text*, pag. 7

Indeed, to estimate these parameters the method of maximum likelihood is applied, which allows to find the values of α and βs that maximize the probability of obtaining the data set.

However, it is easier to work with the log-likelihood, expressed as

$$l = \log(L) = \sum_{i=1}^{n} [Y^i \log \mathrm{G}(\beta^i X^i) + (1 - Y^i)\log(1 - G(\beta^i X^i))],$$

from which the first order derivative is calculated and then solved for α and βs

Since the first order derivative is nonlinear and non-analytic, numerical optimization methods can be applied, involving reiteration through a software.

Besides, it is important to notice that, in order to perform a study based on this model, there are some conditions that need to be respected:

- The dependent variable of the logistic regression has to be discrete and mostly dichotomous;
- The dependent variable needs to be coded accordingly to the event occurring, meaning that the outcome wished should be coded to be 1;
- Observations have to be independent and multicollinearity not present, i.e. independent variables which are not linear functions of each other;
- The model should not be under fitted with meaningful variable included, or overfitted with meaningless ones. It must be correctly fitted;

- Dependent and independent variables could not be in a linear relationship, but independent ones are linearly related to the log odds of an event[21];
- Since maximum likelihood has lower power to estimate parameters than ordinary least squares, the sample size need to be large.

Nevertheless, this model generally is used for only cross-sectional data, which do not involve the time dimension, because ignored or not available, and, for the purpose of this work where time enters in the analysis, a model involving also this sphere is more appropriate. This will be presented in the following section.

3.3.2 Survival Analysis, Cox Proportional Hazards Model and Extended Version for Time-Dependent Variables

Survival data is a popular data analysis approach, consisting in a collection of statistical procedures whose outcome variable studied is the period before an event occurs.

The variable time, expressed in days, weeks, monthlies or years, is referred as survival time, since it provides the time an individual "survived" during the follow-up period, i.e. the period explored, before an event occurs.

The event, instead, is defined as failure, since generally it refers to a negative outcome for the individual.

[21] Odds of an event are the ratio of the probability that an event will occur to the probability that the same event will not.

The goal of survival analysis is to estimate the distribution function

$$S(t) = P(T > t) = \int_t^{\infty} f(u)du = 1 - F(t)$$

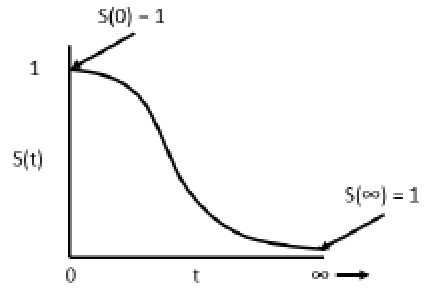
$$h(t) = \lim_{\Delta \to 0} \left(\frac{P(t \leq T \leq t + \Delta t | T \geq t}{\Delta t} \right) = \frac{f(t)}{S(t)}$$

where $S(t)$ and $h(t)$ are respectively the survival function and the hazard ratio, denoting T the random variable for survival time of the individual and t any specific value of that random variable.

The survivor function $S(t)$ provides the probability that a person a person survives longer than a specific time t, and theoretically is characterised by the following properties, reflected also in graph n° 9:

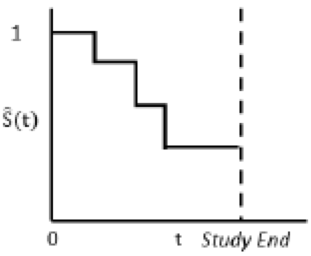
- it can be represented as a smoothed nonincreasing curve;
- at time $t = 0$, $S(t) = 0 = 1$, meaning that the probability of surviving past time 0 is 1, since no one was subject to the event yet at the beginning of the period;
- at time $t = \infty$, $S(t) = \infty = 0$, meaning that the survivor curve will fall to 0 if the period increased without any limit since nobody would survive.

Graph n° 9[22]



Nevertheless, when actual data are used, in reality, graphs are not a smoothed curve but more a steps functions, as illustrated in graph n° 10

Graph n° 10[23]



The hazard function $h(t)$, instead, gives the instantaneous failure rate of individuals at time t conditional on survival until t. Since it focuses on failing, in contrast to the survivor function which focuses on not failing, it can be considered as a dealer of opposite information respect the survivor curve.

[22] M.Billio, M. Riedel, A. Bedin, D. Leboullenger, *Technical Report on the Methodology Design to Carry Out Portfolio Analysis,* L. Bertalot, pag. 23

[23] *Ibidem*

However, there is a relationship between the two that can be summarised as:

$$S(t) = \exp[-\int_0^t h(u)du]$$

and

$$h(t) = -\left[\frac{dS(t)/dt}{S(t)}\right]$$

Even if the survival analysis is well used, two analytical problems may arise, and attention need to be put on that: censoring and truncation.

As far as the first one, the term censoring, and more precisely right-censoring, indicates the situation in which a subject is recorded in the dataset for the study, but information about his survival period is not available.

There are three reasons which can bring to censoring:

- quickly end of the study;
- lost to follow up, i.e. the individual under analysis disappears for unknow reasons;
- withdrawals.

Not accounting for censoring will lead to unbiased results.

Indeed, the likelihood function for the sample will change depending if censoring happens or not. The likelihood function will pass from being

$$L\{\beta|(t_f^i, X^i), \dots, (t_f^n, X^n)\} = \prod_{i=1}^{n} S(t_f^i|\beta, X^i)h(t_f^i|\beta, X^i)$$

to being

$$L\{\beta|(t_f^i, X^i), \dots, (t_f^n, X^n)\} = \prod_{i=1}^{n} S(t_f^i\,|\beta, X^i)$$

A way, to take censoring into account, is to produce a dichotomous variable δ, assuming different value depending of censoring happened or not:

$$\delta = \begin{cases} 0, if \quad x < t_{RC} \ (uncensored) \\ 1, if \quad x \geq t_{RC} \ (censored) \end{cases}$$

where the observed lifetime is $Z = \min(t_f, t_{RC})$, with t_f indicating the failure time and t_{RC} the censoring time. In this way, the individual is randomly censored if censoring happens prior to the occurrence of the event.

Besides, when dealing with fully paid off loans, the last payment date recorded represents the censoring time, since the idea is that that despite these loans not defaulted during the specific contractual period, they could have not survived in the future if the period had got longer. Due to this reason, the life time is generally expressed as (Z^i, δ^i).

Regarding truncation, instead, and most precisely left-truncation, it happens when an individual is at risk prior to the analysis beginning. Taking again loans as examples, they may have been created before the study has started but due to not availability of information data are reported from the time in which the study has begun.

In this case, as in the previous one, left truncation can lead to unbiased results.

Two type of left-truncation exists, if a loan is created at $t_0 < t_{LT}$, where t_{LT} indicates the moment in which the studies begins effectively:

- $t_f < t_{LT}$, when the default happens before the study starts and so the loan is not included in the analysis;
- $t_f > t_{LT}$, when the loan survives enough to take part to the analysis, but defaults before the ending type.

Also in this case, the likelihood function will change, resulting in

$$L^i = \{\beta|(t_f^i, X^i)\} = \left[\frac{s(t_f^i|\beta, X^i)}{s(t_{fLT}|\beta, X^i)}\right] h(t_f^i|\beta, X^i),$$

where the ratio $\frac{s(t_f^i|\beta, X^i)}{s(t_{fLT}|\beta, X^i)}$ is the probability of surviving until t_f^i given survival up to t_{fLT}.

Now that a general overview of survival analysis, with its limits, has been presented, it's time to introduce its frequent model, the Cox Proportional Hazards Model.

It is expressed in term of the hazard model formula, as the product of two terms:

$$h(t, X) = h_0(t)e^{\sum_{i=1}^{p} \beta_l X_l}$$

where $h_0(t)$ is the baseline hazard function at time t and $e^{\sum_{i=1}^{p} \beta_l X_l}$ is the exponential expression to the linear sum of $\beta_i X_l$, with p the number of covariates X and β_i is the parameter that has to be estimated for ith covariate. The first term is called baseline function, because if all the X are equal to 0, the formula reduces to $h_0(t)$.

Actually, an important characteristic is that the baseline function is a function of t only and does not depend on covariates X, while the exponential expressions involves only covariates X and not t. In this way, X are time independent.

Moreover, since the baseline hazard is an unspecified function, the Cox model is generally called a semiparametric model. In this way is not feasible to estimate the β coefficients, but the hazard ratio can be used, to overcome this problem.

The hazard ration (HR) is expressed as the hazard of one subject in the study over the hazard of another one, which differentiate for the values of their respective covariates.

If t_k^i and t_k^j refer to observations time of individual i and j respectively and X^i X^j their covariates, the hazard ratio can be written in estimate term as:

$$\widehat{HR} = \frac{\hat{h}(t_k^i, X^i)}{\hat{h}(t_k^j, X^j)}$$

It is possible now to substitute the Cox model formula in the previous one, obtaining an expression for estimating βs.

$$\widehat{HR} = \frac{\hat{h}(t_k^i, X^i)}{\hat{h}(t_k^j, X^j)} = \frac{\widehat{h_0}(t_k^i)e^{\sum_{l=1}^{p}\widehat{\beta_l}x_l^i}}{\widehat{h_0}(t_k^j e^{\sum_{l=1}^{p}\widehat{\beta_l}x_l^j})} = e^{\sum_{l=1}^{p}\beta_l(x_l^i x_l^j)} = \theta,$$

where θ is a time-independent constant.

Unfortunately, in an analysis like the one that will be performed in this study involving loan information, time represents an important part, since several variables are prone to change across time.

The extended version of the Cox model allows for this:

$$h\big(t, X(t)\big) = h_0(t)\exp\left[\sum_{l=1}^{p=1}\beta_l x_l + \sum_{l=1}^{p=2}\delta_l x_l\,(t)\right],$$

What changes from the previous one is that the exponential part of the equation contains p_1 time-independent predictors, denoted by x_l and p_2 time-dependent predictors denoted by $x_l(t)$.

The HR for the extended Cox model version is given by:

$$\widehat{HR} = \frac{\hat{h}(t_k^i, X^i)}{\hat{h}(t_k^j, X^j)} = \exp\left[\sum_{l=1}^{p_1}\beta_l\big(x_l^i - x_l^j\big) + \sum_{l=1}^{p_2}\delta_l[x_l^i(t) - x_l^i(t)]\right],$$

where $x_l^i(t)$ and $x_l^i(t)$ indicate the specifications at time t for the set of predictors containing both time-independent and time-dependent variables.

4. Data Description

The purpose of this chapter is to present the construction of the data set on which the analysis is based. Firstly, the procedure regarding the selection and the aggregation of loan-level data is presented, secondly the definition of energy efficient building ratings and the energy initiatives proposed is supplied and thirdly the choice of variables applied for models and summary statistic is described.

4.1 Construction of the Sample

In the analysis performed, data regarding loan-level information has been taken from European DataWarehouse (ED), which provides a periodically updated dataset of securitized European mortgages. It represents part of the European Central Bank ABS Loan Level Initiative, established in 2012, and gives an open platform for users to access over ABS data transactions and private portfolios from different originators in Europe.

Moreover, European Central Bank's website provides a comprehensive overview of loan-level data templates with detailed descriptions of variables regarding residential mortgages-backed securities (RMBS).

The asset country chosen for the analysis is Italy, with a period ranging from to March 2014 to October 2018.

Regarding houses selected, the analysis is restricted to residential houses, in particular “detached/semi-detached”, “apartment” or “terraced house” (respectively AR131=1,2, or 3 where AR131 stands for property type), with construction year (AR133) ranging from 1900 to 2018, whose occupancy type is restricted to “owner-occupied” (AR130=1).

As far as repurchased loans, they are eliminated from the sample (AR166=1 to 4).

Moreover, only individual borrowers are selected, with each one associated to only one building and vice versa, whose total income is given by the sum of primary one and secondary one, if the first is different from 0 (AR26 and AR28 respectively).

Final dataset amounts to 125,321 mortgages.

After having applied this selection to the original database, two other subsamples have been created, with data relating one to Lombardia and the other one to Trentino Alto Adige. In this case the number of mortgages to deal is restricted respectively to 31,141 and 663. For Lombardia, the province of Monza is not considered for lack of data, while for Trentino Alto Adige, the sample is obtained only with data from the province of Trento, since no data available for Bolzano in terms of energy variables.

4.2 Energy Efficiency Choice

4.2.1 Italy

Whereas the analysis focuses on the relation between the probability of default and energy efficiency, attention is needed on how the second one is considered.

A proxy has been employed to enter energy efficiency in the model regarding Italy.

Data related to energy has been taken by ENEA, i.e. Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico, which is the Italian energy agency providing the amount of annual fiscal detraction for energy efficiency improvement granted between 2014 and 2017, standardized among the total numbers of properties in each region.

This number can be seen as a proxy for government investments in each region, since better examples are not available.

4.2.2 Lombardia and Trentino Alto Adige

Making the analysis more specific, focusing only on single regions, information coming from "APE" certificate (Attestato di Certificazione Energetica) can be exploited.

This document gives an overview of property's energy performance in terms of annual amount of primary energy consumed with a standard use of the property, expressed by one or more descriptors considering levels of insulation and technical installation characteristics, the various energy needs and for tertiary sector, lighting, escalators and lifts.

The obligation of this certificate exists when dealing with:

- new construction and extensions;
- major first/second level renovations;
- buildings used by the Public Administration and open to the public with a usable area of more than 250 m^2;

- new or renewed contracts Energy Service and Energy Plus;
- new or renewed contracts for management of heating systems;
- new or renewed lease contracts.

Among the possible variables obtainable for energy efficiency, this analysis will deal with 4 of them:

- ETH (Thermal Energy-hot), which is the ideal amount of thermal energy required by the building for winter air conditioning, during the heating season;
- ETC (Thermal Energy-cold), which is the ideal amount of thermal energy required by the building for summer air conditioning, during the cooling season;
- EPW (Energy requirements-water), which is the ideal amount of thermal energy required by the building for the production of domestic hot water;

For each one of these, a mean value is obtained dividing the number of energy related to one province, obtained by summing up energy amounts relative to each dwellings, to the number of homes in that province.

4.3 Choice of Variables

After having presented how the data set is built and the current energy initiatives, it is important to have a look to the variables that entered in the analysis. The dependent variable is a dummy indicating whether a borrower is in default (1) or not (0), considering a mortgage in default when in arrears for more than 90 days, according to Basel committee definition.

For independent variables, instead, past literature has been taken as example, indicating which variables have a significant impact on the probability of default of mortgages. These variables can be subdivided into four different types: mortgage, building, borrower and macroeconomic/financial variables.

Regarding mortgage class, the variables employed are loan-to-value ratio (LTV), debt to income ratio (DTI), loan term and debt service coverage ratio (DSCR).

The LTV, which is a proxy used by financial institutions to asses lending risk, is computed by dividing the origination/current loan balance (AR66/AR67) by the valuation amount of the property involved (AR136).

The DTI, which is defined as the percentage of gross monthly income available for monthly debt payments, is generated by dividing origination/current loan balance per building by total household income available.

Loan term, which is given by the difference between issuance and maturity date, measured in month, is provided directly by the original dataset (AR61).

DSCR, which represents a value-weighted monthly periodic payment for each component on the same building, is obtained by using the current loan balance, interest rate (AR109) and the number of periods left until maturity, given by the difference between the date of loan maturity (AR56) and the poolcutoffdate.

As far as building class, property type (AR131) is considered.

Concerning borrowers, instead, the variables related are total income, defined as the sum of primary and secondary income (AR26, AR28), and borrower age at origination, given by the difference between the age of loan origination (AR55) and the borrower year of birth (AR18).

Finally, to control for the overall economic situation, macroeconomic variables have been included: the Italian unemployment rate, the end of month 10-year German government bond yields, the monthly standard deviation of the 10-year German bond yields and the yield curve slope defined as the difference between 10- and 1-year EUR swap rates, all taken from Bloomberg.

4.3.1 Summary Statistics

In the following, summary statistics at property level are reported. They represent cross-sectional one-time snapshot of the main mortgage, building and, borrower characteristics using the latest reported values, differentiating for defaulted and non-defaulted loans. Tables n° 5,6,7 provide information respectively to Italy, Lombardia and Lombardia plus Trentino.

Table n° 5

Variable	Default	Mean	Median	SD	Min	Max	N
Borrower age	0	40	39	9	22	66	123,673
	1	39	38	9	22	66	1,648
Borrower income, total	0	30,578.40	24,000.00	18,699.02	2,000.00	203,853.30	125,421
	1	27,677.25	23,334.72	16,553.29	3,265.60	147,667.73	1,672
Current debt to income	0	3.42	3.23	1.95	0.01	16.16	125,421
	1	3.90	3.64	2.20	0.05	15.05	1,672
DSCR, current	0	4.58	3.64	2.29	1.25	14.81	125,421
	1	3.57	3.27	1.55	1.25	13.91	1,672
Interest rate	0	2.19	2.00	1.25	0.00	8.00	125,421
	1	2.22	1.83	1.72	0.00	7.60	1,672
LTV, current	0	0.50	0.53	0.24	0.00	0.89	125,421
	1	0.50	0.53	0.21	0.01	0.89	1,672
LTV, original	0	0.67	0.72	0.21	0.14	1.14	124,187
	1	0.69	0.75	0.18	0.14	1.14	1,650
Month, current	0	8.26	9.00	1.60	2.00	12.00	125,421
	1	6.37	6.00	2.08	2.00	12.00	1,672
Mortgage age (in years)	0	6.38	5.83	4.16	0.00	20.33	125,421
	1	9.43	10.08	3.70	1.50	18.17	1,672
Mortgage term (in years)	0	22.88	25.00	6.50	3.00	42.58	125,421
	1	24.25	25.08	6.19	7.17	40.08	1,672
Original balance, total	0	120,342	112,000	53,927	30,000	370,000	123,951
	1	127,018	120,000	55,721	30,000	370,000	1,636
Original debt to income	0	4.54	4.36	2.03	0.41	36.15	123,951
	1	5.27	4.81	2.35	0.91	17.50	1,636

Taking Italy as a whole, the age of borrowers taking a loan is the same for both defaulted and not defaulted ones. The income, instead, changes a little between the two categories, with borrowers having an average amount of € 30,578.40 , while defaulting borrowers having an income of € 27,677.25, which is a plausible result. Original balance changes from € 120,342 to € 127,018, while interest rates are similar.

Moreover, if single regions are taken in consideration, the distance for these variables between default situation and not default one, is reflected more or less.

Table n° 6

Variable	Default	Mean	Median	SD	Min	Max	N
Borrower age	0	38	37	9	22	66	30,753
	1	38	37	9	22	66	388
Borrower income, total	0	32,924.32	26,281.00	19,913.17	2,000.00	182,977.20	30,753
	1	28,862.12	26,086.79	14,431.59	9,055.12	134,717.00	388
Construction Year	0	1986	1992	22	1900	2018	7,952
	1	1986	1995	26	1900	2011	27
Current debt to income	0	3.36	3.13	1.96	0.02	15.76	30,753
	1	3.55	3.48	1.59	0.22	13.47	388
DSCR, current	0	4.78	3.83	2.37	1.25	14.81	30,753
	1	3.82	3.76	1.28	1.36	9.37	388
Interest rate	0	1.97	1.78	1.24	0.00	7.50	30,753
	1	1.73	1.27	1.58	0.00	7.00	388
LTV, current	0	0.52	0.57	0.25	0.00	0.89	30,753
	1	0.52	0.54	0.19	0.02	0.88	388
LTV, original	0	0.71	0.77	0.20	0.14	1.14	30,436
	1	0.73	0.77	0.16	0.16	1.14	380
Month, current	0	8.36	9.00	1.38	2.00	12.00	30,753
	1	7.21	8.00	1.60	2.00	12.00	388
Mortgage age (in years)	0	7.15	6.67	4.55	0.00	20.17	30,753
	1	10.92	11.92	3.63	1.83	18.17	388
Mortgage term (in years)	0	23.97	25.08	6.31	4.58	42.58	30,753
	1	25.72	25.42	5.24	10.08	40.00	388
Original balance, total	0	128,558	120,000	54,181	30,000	370,000	30,465
	1	133,284	125,000	52,561	46,481	370,000	385
Original debt to income	0	4.49	4.45	1.71	1.18	10.92	29,884
	1	4.87	4.75	1.37	1.23	10.58	382

Table n° 7

Variable	Default	Mean	Median	SD	Min	Max	N
Borrower age	0	38	37	9	22	66	31,385
	1	38	37	9	22	66	392
Borrower income, total	0	32,823.51	26,193.00	19,872.40	2,000.00	182,977.20	31,385
	1	28,891.66	26,086.79	14,393.64	9,055.12	134,717.00	392
Current debt to income	0	3.37	3.15	1.96	0.02	15.76	31,385
	1	3.55	3.48	1.59	0.22	13.47	392
DSCR, current	0	4.76	3.82	2.36	1.25	14.81	31,385
	1	3.81	3.75	1.28	1.36	9.37	392
Interest rate	0	1.97	1.75	1.23	0.00	7.50	31,385
	1	1.74	1.27	1.59	0.00	7.00	392
LTV, current	0	0.52	0.56	0.25	0.00	0.89	31,385
	1	0.52	0.54	0.19	0.02	0.88	392
LTV, original	0	0.71	0.76	0.20	0.14	1.14	31,053
	1	0.73	0.77	0.16	0.16	1.14	384
Month, current	0	8.34	9.00	1.41	2.00	12.00	31,385
	1	7.21	8.00	1.60	2.00	12.00	392
Mortgage age (in years)	0	7.10	6.58	4.55	0.00	20.17	31,385
	1	10.91	11.92	3.64	1.83	18.17	392
Mortgage term (in years)	0	23.91	25.00	6.33	3.67	42.58	31,385
	1	25.66	25.42	5.28	10.08	40.00	392
Original balance, total	0	128,538	120,000	54,284	30,000	370,000	31,085
	1	133,552	125,000	52,818	46,481	370,000	389
Original debt to income	0	4.54	4.45	1.92	0.48	29.45	31,085
	1	4.94	4.76	1.56	1.23	14.71	389

5. Empirical Results

5.1 Italy

As previously defined, for the analysis of the whole Italian market, the model is applied to a cross-section of loan-level data taken from European DataWarehouse, covering the period 2014 to 2018. For each loan, the most recent observation in the sample are considered.

Moreover, the amount of annual fiscal detraction for energy efficiency improvement granted between 2014 and 2017, standardized among the total numbers of properties in each region, is used as proxy for energy efficiency level. For visual reasons, each decimal number is multiplied by 1000.

5.1.1 Logistic Regression

Firstly, a logistic regression is applied. Table n° 8 reports the odds ratio obtained[24].

[24] The odds ratio is expressed as a number ranging from zero, meaning that the event will never happen, to infinity, where the event is certain to occur. The value of one is used as reference point; an odds ratio lower than one is evidence of the positive association between the two variables, while in the opposite case a negative association is supposed

Table n° 8[25]

	(1)	(2)	(3)	(4)
EE investments	0.759787***	0.6943414***	0.7464268***	0.7464268***
	[0.0262052]	[0.0239754]	[0.02778]	[0.0526207]
Current LTV		0.4628646***	1.228011	1.228011
		[0.0722505]	[0.224148]	[0.2383848]
Current DTI		0.8035895***	0.7212784***	0.7212784***
		[0.0241066]	[0.0231356]	[0.0416668]
DSCR		0.5417431***	0.5580177***	0.5580177***
		[0.0222954]	[0.02251]	[0.347449]
Mortgage Term		11.41328***	11.50775***	11.50775***
		[1.938591]	[1.116607]	[3.604579]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Region Cl.
Observations	125,321	125,321	125,321	125,321
Pseudo R-squared	0.0036	0.0585	0.1171	0.1171

The first column shows the outcome with only EE investments as independent variable, without taking in consideration dwellings, households and mortgage side, and neither macroeconomic factors.

The odds ratio of 0.759787 demonstrates how the probability of default diminishes in region with higher EE investments, with a significant level of 1%.

Nevertheless, many factors have to be considered for a complete analysis.

The second column shows the estimation taking in consideration also dwellings characteristics, in terms of building type, households' characteristics, in terms of

[25] The statistical significance of 10%, 5%, 1% is denoted respectively by ***,**,*.

borrower's age at origination and total income, and mortgage characteristics, in terms of current LTV, current DTI, DSCR and mortgage terms. The association between the probability of default and the energy efficient proxy remains still negative and significant.

In the third column, the result is obtained by a regression which includes, beyond the factors just described, also macroeconomic covariates in terms of unemployment, government bond yields, volatility of government bond yields and yield curve slope. It confirms the considerations made until now.

While these three specifications have robust standard errors, the last column, instead, provides the outcome when standard errors are clustered at regional level, including all the possible factors, but this doesn't affect the findings reported in the third specification.

Findings can be validated using 3 robustness checks, with results showed in table n° 9, which includes only the estimations and standard errors of the energy efficient variable. Here the baseline model is the one represented in column four of table n° 8, to which changes are applied.

Table n° 9

Model	OR(EE)	Std.
Spec 1	0.7858991***	[0.523545]
Spec 2	0.7447746***	[0.588536]
Spec 3	0.7475255***	[0.0526929]

In column 1, current LTV and DTI are substituted with the original ones, since it is common to estimate a credit risk model with original covariates. In column 2, the real borrower age at origination is used, while in column 3 it is categorized in 10 subsamples. The redefinition of variables does not affect the overall relationship between borrower's probability of default and EE investments.

5.1.2 Cox Hazard Model

The problem with the logistic regression, as previously mentioned, comes from the lack of consideration for the time dimension, and the for this reason the Cox Hazard model is applied, especially in its extended version, which allows for time-dependent variables.

Table n° 10 provides the estimate when this model is applied.

Table n° 10

	(1)	(2)	(3)	(4)
EE investments	0.7758977***	0.7147429***	0.7676805***	0.7676805***
	[0.02658]	[0.0243152]	[0.0286034]	[0.0626825]
Current LTV		0.9995386***	1.00014	1.00014
		[0.0000738]	[0.0000899]	[0.0000949]
Current DTI		0.9998801***	0.9998431***	0.9998431***
		[0.0000145]	[0.000015]	[0.0000268]
DSCR		0.9996901***	0.999726***	0.999726***
		[0.0000198]	[0.0000188]	[0.0000291]
Mortgage Term		1.001166***	1.001181***	1.001181***
		[0.0000785]	[0.0000847]	[00.0001492]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Region Cl.
Observations	125,321	125,321	125,321	125,321

The four regressions follow the scheme of the logistic case.

For the first regression, allowing for only the energy efficient variable, the estimate confirms significatively the negative relationship between default and energy variable, with a coefficient of 0.7758977.

For the second, in opposite with the logistic case, the time dimension is considered, with current LTV, current DTI, DSCR, loan term and total income treated as time-varying covariates.

LTV and DSCR varies across time, as the loan components are being repaired, while loan term, total income and consequently DTI varies less often.

Loan term is affected by any additional loan components that are added to the already existed ones; total income changes according to variation in the income status or salary of borrowers.

However, also in this case there is evidence of negative association, with a significance level of 1%.

In the third regression, beyond time-varying covariates related to borrower and mortgage characteristics, also macroeconomic variables are seen as time-dependent, resulting in a negative relationship as well; evidence confirmed also in the fourth regression allowing for errors clustered at provincial level.

Robustness of results is checked in table n° 11

Table n° 11

Model	OR(EE)	Std.
Spec 2	0.7664511***	[0.0623884]
Spec 3	0.7684424***	[0.0627219]

In this case, specification n° 1, substituting current LTV and DTI with original ones, is omitted as the main property of Cox regression is to include original as well as current covariate values in the regression analysis. Specification n° 2 and 3 replace the borrower age with real one and age categorized in 10 years respectively.

The Cox Model confirms the result obtained in the logistic case.

5.2 Lombardia

After considering Italy in its entirety, the analysis concentrates on single regions. The reason is to understand whether, and to what extent, the provincial energy efficient level impacts on the mortgages' default rates in the considered province.

Starting from Lombardia (excluding for Monza province due to lack of data), the variables used as proxy for energy efficient, as already said, are ETH, which is the ideal amount of thermal energy required by the building for winter air conditioning, during the heating season, ETC, which is the ideal amount of thermal energy required by the building for summer air conditioning, during the cooling season and EPW, which is the ideal amount of thermal energy required by the building for the production of domestic hot water.

5.2.1 Logistic Regression

As far as ETH, the results of logistic regressions are presented in table n° 12.

Table n° 12

	(1)	(2)	(3)	(4)
ETH	1.118149***	1.117388***	1.093761**	1.093761*
	[0.0373581]	[0.0383242]	[0.387963]	[0.0587905]
Current LTV		0.4131366**	0.7968678	0.7968678
		[0.1485484]	[0.3162917]	[0.3688134]
Current DTI		0.6242477***	0.531317***	0.531317***
		[0.0410387]	[0.035342]	[0.0381508]
DSCR		0.4921862***	0.472873***	0.472873***
		[0.035021]	[0.033382]	[0.0565141]
Mortgage Term		79.40311***	118.1913***	118.1913***
		[32.85748]	[49.37484]	[44.16448]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	30,506	30,056	30,506
Pseudo R-squared	0.0026	0.0737	0.1069	0.1069

In the first column, the outcome is obtained considering ETH only. In this case, the relationship between the energy efficient value and the probability of default is positive, since the odds ratio amounts to 1.118149 but the conclusions are the same of the Italian case. In fact, ETH is a measure of energy consumption in dwellings, and the higher ETH, the lower the energy efficiency (EE). Indeed, as ETH increases, and so EE decreases, probability of default of borrowers increases.

However, as in the previous analysis, other factors have to be added, with second specification taking mortgages, households, dwelling characteristics, and third specification taking, over these, macroeconomics characteristics, both using robust standard errors. The positive relationship between probability of default and ETH remains valid and the estimates result statistically significant at 1% level and 5% respectively.

When standard errors are clustered at provincial level, instead, as in column four, the level of significance reduces to 10%.

The robustness checks are presented in table n° 13, where fourth regression is changed with specification 1 substituting for original LTV and DTI the respective current variables, specification 2 replacing with the real borrower age at origination the inherent variable and specification 3 with real borrower age categorized for 10 years.

Table n° 13

Model	OR(EE)	Std.
Spec 1	1.078286	[0.0548807]
Spec 2	1.093734*	[0.0574315]
Spec 3	1.095443*	[0.589319]

The positive relationship is confirmed at a significant level of 10% for specification 2 and 3, with substitution not affecting the results. For specification 1, instead, the estimate turns out to be not significant.

The same structure is repeated for ETC and EPW, but the results are not the one hoped. Tables n° 14 and 15 provide the outcomes considering ETC and its specification, while tables n° 16 and 17 give the outcomes for EPW.

Table n° 14

	(1)	(2)	(3)	(4)
ETC	0.9697733***	0.9969864***	0.9710154***	0.9710154**
	[0.0077593]	[0.0076441]	[0.0082752]	[0.0126229]
Current LTV		0.40664375**	0.7879567	0.7879567
		[0.1457754]	[0.3127309]	[0.3607198]
Current DTI		0.6192041***	0.5274215***	0.5274215***
		[0.0407312]	[0.3051464]	[0.038985]
DSCR		0.4871801***	0.4692398***	0.4692398***
		[0.0347298]	[0.332116]	[0.567185]
Mortgage Term		82.96138***	123.9768***	123.9768***
		[34.48473]	[52.16812]	[48.34085]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	30,506	30,506	30,506
Pseudo R-squared	0.0029	0.0748	0.1078	0.1078

Table n° 15

Model	OR(EE)	Std.
Spec 1	0.9761385*	[0.104044]
Spec 2	0.9711145**	[0.012391]
Spec 3	0.9698453**	[0.0127248]

Table n° 16

	(1)	(2)	(3)	(4)
EPW	0.9984107***	0.99808***	0.9983668***	0.9983668***
	[0.0003406]	[0.0003463]	[0.000374]	[0.0003131]
Current LTV		0.3981731***	0.79025454	0.79025454
		[0.1431805]	[0.314596]	[0.4001172]
Current DTI		0.6152187***	0.5264847***	0.5264847***
		[0.0403238]	[0.0352621]	[0.0379377]
DSCR		0.4476544***	0.4641151***	0.4641151***
		[0.343733]	[0.0332041]	[0.0559234]
Mortgage Term		88.22687***	130.8001***	130.8001***
		[36.81536]	[55.43526]	[48.45044]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	30,506	30,506	30,506
Pseudo R-squared	0.0051	0.0784	0.1102	0.1102

Table n° 17

Model	OR(EE)	Std.
Spec 1	0.9983249***	[0.000385]
Spec 2	0.9983551***	[0.0003114]
Spec 3	0.9983274***	[0.003108]

In both cases, the association between probability of default and the related energy variables turn out to be negative, with probability of default increasing with the energy efficiency increasing as well.

This incoherence of results has a reason behind: even though the Lombardia region provides a very detailed dataset on dwellings' energy consumption information, its potential cannot be fully exploited, due to the fact that it is not possible to merge dwellings' energy consumption information one-to-one with the anonymized loan-level data, because the last subset, for data privacy regulation, does not supply the number or street names of dwellings to which it refers.

However, this is a superable obstacle, if banks at the moment of loan originations, would collect buildings' energy efficient information.

5.2.2 Hazard Cox Model

The outcomes obtained for ETH variable, reperforming the analysis with the application of the Hazard Cox Model, are shown in tables n° 18 and 19.

Table n° 18

	(1)	(2)	(3)	(4)
ETH	1.107993***	1.104002***	1.083031**	1.083031
	[0.0363357]	[0.0369592]	[0.0368762]	[0.058257]
Current LTV		0.9995522***	0.9999137	0.9999137
		[0.000167]	[0.0001898]	[0.0002206]
Current DTI		0.9997638***	0.9996992***	0.9996992***
		[0.0000304]	[0.0000305]	[0.0000336]
DSCR		0.9996517***	0.9996494***	0.9996494***
		[0.0000322]	[0.0000318]	[0.0000518]
Mortgage Term		1.001949***	1.002284***	1.002284***
		[0.000186]	[0.0001978]	[0.0001922]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	31,141	31,141	31,141

The regressions, with robust standard errors, confirm the significant positive association between the probability of default and the energy related variable, but once clustered errors are applied the estimate becomes not significant. Also the robustness checks are not confirmed in table n° 19.

Table n° 19

Model	OR(EE)	Std.
Spec 2	1.084115	[0.0569201]
Spec 3	1.08527	[0.0586212]

Tables n° 20, 21, 22, 23 reports the results for regressions and robustness check sfor ETC variable and EPW one, confirming the considerations made for the logistic case.

Table n° 20

	(1)	(2)	(3)	(4)
ETC	0.9733939***	0.9703962***	0.9714727***	0.9714727**
	[0.0077416]	[0.0076054]	[0.0079418]	[0.0129927]
Current LTV		0.9995459***	0.9999037	0.9999037
		[0.0001667]	[0.00019]	[0.0002213]
Current DTI		0.9997608***	0.9996958***	0.9996958***
		[0.0000304]	[0.0000306]	[0.0000327]
DSCR		0.9996475***	0.9996454***	0.9996454***
		[0.0000322]	[0.0000319]	[0.0000528]
Mortgage Term		1.001967***	1.002305***	1.002305***
		[0.0001867]	[0.0001987]	[0.0001993]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	31,141	31,141	31,141

Table n° 21

Model	OR(EE)	Std.
Spec 2	0.9714188**	[0.0125903]
Spec 3	0.970267**	[0.0132613]

Table n° 22

	(1)	(2)	(3)	(4)
EPW	0.9985442***	0.9983483***	0.9983934***	0.9983934***
	[0.0003373]	[0.0003373]	[0.0003584]	[0.0003584]
Current LTV		0.9995419***	0.9999123	0.9999123
		[0.000167]	[0.0001908]	[0.0001908
Current DTI		0.9997603***	0.9996952***	0.9996952***
		[0.0000302]	[0.0000308]	[0.0000308]
DSCR		0.9996413***	0.9996404***	0.9996404***
		[0.0000323]	[0.0000322]	[0.0000322]
Mortgage Term		1.001988***	1.002328***	1.002328***
		[0.0001868]	[0.0001985]	[0.0001985]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,141	31,141	31,141	31,141

Table n° 23

Model	OR(EE)	Std.
Spec 2	0.9983679***	[0.0002695]
Spec 3	0.9983452***	[0.0002786]

5.3 Lombardia and Trentino Alto Adige

After having concentrated on only one region, it is interesting to see what is the effect when two neighbouring regions are considered, and if the addition of the second one does influence positively or negatively the result.

The region chosen for this type of analysis was Trentino Alto Adige, but with energy data related only to the Trento province, since Bolzano ones where not available.

Indeed, a new dataset was created with value referring to both Lombardia and the only province of Trento. The energetic variable taken in consideration, in this case, where only ETH and EPW, since ETC was not available for Trento.

5.3.1 Logistic Regression

The steps made for the logistic regression in the Lombardia case are repeated considering also Trento, starting firstly with energy variable as the only independent variable, secondly adding mortgage, borrower and dwellings related variables, and thirdly and quarterly adding also macroeconomic variables with robust errors and clustered ones respectively.

Table n° 24 reports the outcomes for the ETH variable, and table n° 25 the odds ratio when changes to the fourth specification are made (original LTV and DTI, real borrower age and borrower age categorized for 10 years).

Table n° 24

	(1)	(2)	(3)	(4)
ETH	1.106555***	1.109813***	1.108191**	1.08191
	[0.0370709]	[0.0377394]	[0.0385233]	[0.0586765]
Current LTV		0.4759233**	0.9020596	0.9020596
		[0.1677501]	[0.0334793]	[0.3947696
Current DTI		0.0390383***	0.5180205***	0.5180205***
		[0.3090383]	[0.0334793]	[0.348223]
DSCR		0.0345463***	0.4628504***	0.4628504***
		[0.0345463]	[0.0324552]	[0.538454]
Mortgage Term		78.85675***	121.3178***	121.3178***
		[32.29347]	[49.89253]	[44.06201]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,777	31,142	31,142	31142
Pseudo R-squared	0.0021	0.0744	0.1070	0.107

Table n° 25

Model	OR(EE)	Std.
Spec 1	1.061548	[0.533202]
Spec 2	1.081976	[0.0568952]
Spec 3	1.084371	[0.0588529]

In table n° 24, the estimates shown for the energetic variable are not so different from the case of Lombardia alone (they are slightly lower) and the degree of significance is the same, except for the fourth specification, with the estimate losing the significance. For check of robustness, all the estimates turn out to be not significant, concluding that adding a region, or more precisely a province, do not improve the results, but has a negative impact on them.

In table n° 26 and 27 the same model is reperformed for EPW, obtaining values almost equal to Lombardia alone.

Table n° 26

	(1)	(2)	(3)	(4)
EPW	0.9986494***	0.9982561***	0.9986237***	0.9986237***
	[0.0003179]	[0.000328]	[0.0003562]	[0.000379]
Current LTV		0.4789044**	0.9190619	0.9190619
		[0.1681609]	[0.3586686]	[0.4362974]
Current DTI		0.5999732***	0.5118167***	0.5118167***
		[0.0381524]	[0.0332892]	[0.03047097]
DSCR		0.4713188***	0.4549057***	0.4549057***
		[0.0339251]	[0.32278]	[0.053178]
Mortgage Term		88.10036***	134.2282***	134.2282***
		[36.2804]	[55.92623]	[0.2247464]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,777	31,142	31,142	31,142
Pseudo R-squared	0.0039	0.0785	0.1094	0.1094

Table n° 27

Model	OR(EE)	Std.
Spec 1	0.9986791***	[0.0004586]
Spec 2	0.9986161***	[0.0003799]
Spec 3	0.9985764***	[0.0003757]

Therefore, the reflections possible are the same of paragraph 5.2.1., with conflicting results among energetic variables due to the impossibility of perfectly merge loan-level information and energy information datasets.

5.3.2 Hazard Cox Model

When applying Hazard Cox model, the situation does not get better, confirming significance positive relationship for ETH variable, except for regression four, in table n° 28 and specifications in table n° 29, and negative significant relationship for EPW, confirming the incoherence of results.

Adding the province of Trento does not affect qualitatively the Lombardia situation.

Table n° 28

	(1)	(2)	(3)	(4)
ETH	1.095309***	1.094703***	1.066157*	1.066157
	[0.0360732]	[0.0363702]	[0.0370535]	[0.0590534]
Current LTV		0.9996221**	1.000005	1.000005
		[0.0001636]	[0.0001857]	[0.0002132]
Current DTI		0.9997536***	0.9996831***	0.9996831***
		[0.0000294]	[0.0000298]	[0.000032]
DSCR		0.9996445***	0.9996831***	0.9996831***
		[0.0000323]	[0.0000317]	[0.0000507]
Mortgage Term		1.001953***	1.002305***	1.002305***
		[0.0001837]	[0.000195]	[0.0001859]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,777	31,777	31,777	31,777

Table n° 29

Model	OR(EE)	Std.
Spec 2	1.067555	[0.0571904]
Spec 3	1.068718	[0.0593665]

Table n° 30

	(1)	(2)	(3)	(4)
EPW	0.9987987***	0.9985383***	0.9987226***	0.9987226***
	[0.0003134]	[0.000318]	[0.0003478]	[0.0003441]
Current LTV		0.9996294**	1.000016	1.000016
		[0.000163]	[0.0001862]	[0.0002316]
Current DTI		0.999748***	0.9996782***	0.9996782***
		[0.0000293]	[0.0000301]	[0.0000319]
DSCR		0.9996349***	0.9996295***	0.9996295***
		[0.0000323]	[0.0000321]	[0.0000511]
Mortgage Term		1.001994***	1.002344***	1.002344***
		[0.0001841]	[0.0001956]	[0.0001843]
Dwellings controls	No	Yes	Yes	Yes
Household controls	No	Yes	Yes	Yes
Mortgage controls	No	Yes	Yes	Yes
Market controls	No	No	Yes	Yes
SE	Rob.	Rob.	Rob.	Province Cl.
Observations	31,777	31,777	31,777	31,777

Table n° 31

Model	OR(EE)	Std.
Spec 2	0.9987004***	[0.0003412]
Spec 3	0.9986778***	[0.0003502]

Conclusion

The object of this work is to find a relationship between the borrowers' credit risk for mortgages and the energy efficiency level of the related houses. The idea behind the analysis is that more efficient houses would decrease the borrowers' probability of default, since less utility bills would leave some money apart from disposable income to repay debt and to face unexpected events.

To do so, I worked with European Datawarehouse database, which provides loan-level information for Italy, and then I focalised only on the region of Lombardia and the province of Trento. I merged the respective dataset with energy related variables.

I started by providing a brief explanation of the meaning of green finance, an outlook of the current international and Italian situation and the role of banks in promoting it.

Then I concentrated only on mortgages, pointing out their situation on the European and Italian markets and the state of dwellings' markets.

The rest of the work, instead, focused on the empirical analysis, showing the methodologies used, the creation of the dataset and the choices of variables and lastly the results obtained.

Firstly, the analysis was performed for the Italian using the amount of fiscal detraction divided by the number of dwellings in each region as proxy for the energy variable. Both for the logistic regression and the hazard model a negative relationship between probability of default and energy efficiency is demonstrated, confirming the hypothesis made. Problems come when analysing single region, i.e. Lombardia, or two regions, i.e. Lombardia and Trentino.

For ETH, in majority of the case, the negative relationship of probability of default and energy efficiency is confirmed, but for EPW and ETC is not so. This is due to the impossibility to merge dwellings' energy consumption information one-to-one with the anonymized loan-level data, because the last subset, for data privacy regulation, does not supply the number or street names of dwellings to which it refers.

However, this is a superable obstacle, if banks at the moment of loan originations, would collect buildings' energy efficient information.

The next step in the analysis would be to demonstrate the economic advantage for banks in offering green mortgages, beyond lower probability of default. Actually, lower probability of default, will result in lower loss given default, which would turn out in lower capital requirements.

European Institutions are discussing the convenience to introduce a 'green supporting factor (GSF) or a 'brown penalty factor' in lending.

GSF will represent an instrument which would lower capital requirements for green mortgages, with the consequences of cheaper loans for banks with respect brown mortgages. Two are the possible main drawbacks of this approach: (1) there is no evidence that green mortgages are less riskier than brown ones, since green investments in novel technologies might not be properly tested or could be surpassed; (2) there is not supporting evidence that saving coming from less capital requirements will lead to a lending increase and so grater advantage for financial institutions and clients.

A valid alternative could be a brown penalty factor which, at the contrary, will penalize brown mortgages in terms of higher interest rates, translating in higher capital requirements for banks.

The difference between the two factors is that GSF will lower capital requirements that where recently enhanced after the financial crisis, while the brown penalty factor will act in the opposite way making safer the financial stability system with respect to the actual situation.

References

- Agenzia Italiana per la cooperazione allo sviluppo, *Obiettivi di sviluppo sostenibile* : https://www.aics.gov.it/home-ita/settori/obiettivi-di-sviluppo sostenibile-sdgs/
- An X., Pivo G., *Default Risk of Securitized Commercial Mortgages: Do Sustainability Property Features Matter?* , 2015
- Baccega T., Bedin A., Billio M., Hristova I., Riedel M., *D5.2 Technical report on the portfolio analysis of banks' loan portfolios,* Billio M. (UNIVE), Pellizon L. (SAFE), 2019
- Baccega T., Bedin A., Billio M., Hristova I., Riedel M., *D5.3 Technical Report on the Econometric Assessment and Results,* Billio M. (UNIVE), Pellizon L. (SAFE), 2019
- Billio M., Riedel M., Bedin A., Leboullenger D., *Technical report on the methodology design to carry out portfolio analysis,* Bertalot L. ,2019
- Brockmann K. L., *Green finance-green banking*, KfW Research, 2017
- Cicuto M., *Building's Energy Efficiency and Residential Mortgage Default Behaviour: Evidence from Italy and Germany, 2018.*
- Eichholtz P., Holtermans R., Kok N., Yonder E., *Environmental Performance and the Cost of Capital: Evidence from Commercial Mortgages and REIT Bonds,* 2015
- EeMAP (Energy efficient Mortgages Action Plan), *Creating an Energy Efficient Mortgage for Europe. Building Assessment Briefing: Italy*, 2017

- EeMAP (Energy efficient Mortgages Action Plan), *Creating an Energy Efficient Mortgage for Europe. A review of the state of play on 'green' finance*, 2018
- European Banking Federation, *Towards A Green Finance Framework,*
- European Investment Bank, Green Finance Committee of China Society for Finance and Banking, The need for a common language in Green Finance. Towards a standard-neutral taxonomy for the environmental use of proceeds, 2017
- European Mortgage Federation, *Hypostat 2017. A Review of Europe's Mortgage and Housing Markets,* 2017
- European Mortgage Federation, *Hypostat 2018. A Review of Europe's Mortgage and Housing Markets,* 2018
- Eurosif, *European SRI Study 2018, 2018 : http://www.eurosif.org/wp-content/uploads/2018/11/European-SRI-2018-Study.pdf*
- *Financing the Future. Report of the Italian National Dialogue on Sustainable Finance*
- Fondazione per lo sviluppo sostenibile, *Green Economy Perspectives. Italy Outlook 2017,* 2017: https://www.fondazionesvilupposostenibile.org/wp-content/uploads/dlm_uploads/2017/07/green-economy-perspectives_italy-outlook-2017.pdf
- Gianinazzi V., Pellizzon L., Plazzi A., *Residential mortgage default and positive equity: Lessons from Europe*

- Global Sustainable Investment Alliance, *2018 Global sustainable Investment Review*
- Hera, *Green Financing Framework,* 2019
- Hoover, K. D., *Causality in Economics and Econometrics. An Entry for the New Palgrave Dictionary of Economics,* Durham, 2006
- http://www.ecb.europea.eu/paym/coll/loanlevel/transmission /html/index.en.html
- https://dati.lombardia.it/Energia/CENED-Certificazione-ENergetica-degli-EDifici/rsg3-xhvk
- https://dati.trentino.it7dataset/attestati-di-prestazione-energetica-trentino/resource/23224901-afdf-40ea-a330-4a52b3c6cd1f
- http://www.pbc.gov.cn/english/130721/3133045/index.html
- Kahleborn, Walter, A. Cochu, I.Georgiev, F. Eisinger, D. Hogg, *Defining "green" in the context of green finance*, Luxembourg Pubblication Office of the European Union 2017
- Kaza N., Quercia R. G., Thian C. Y., Home Energy Efficiency and Mortgage Risk, Chapel Hill, 2014
- Kleinbaum D. G., Klein M., *Logistic Regression. A Self-Learning Test*, Springer, 2010
- Peng C. Y. J., Lee K. L., Ingersoll G. M., *An Introduction to Logistic Regression Analysis and Reporting,* Indiana University 2002

- Park, Hyeoun-Ae, *An Introduction to Logistic Regression: From basic Concepts to Interpretation with Particular Attention to Nursing Domain,* Seoul National University, 2013
- Pivo G., *The Effect of Sustainability Features on Mortgage Default Prediction and Risk in Multifamily Rental Housing,* Tucson, 2013
- Richarson S., *Creating an Energy Efficient Mortgage for Europe. Towards a New Market Standard*, Stanford A., Drinkwater J., 2018
- Riedel M., *Buildings' Energy Efficiency and the Probability of Mortgage Default: The Dutch Case,* 2018
- Romania Green Building Council, *Green Homes & Mortgages. A toolkit for Residential Investors and Developers,* 2017
- *UNEP Inquiry,* 2016
- United Nation Climate Change, *The Paris Agreement: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement*
- Van Lenvern F., Ryan-Collins J., *Adjusting Banks' Capital Requirements in Line With Sustainable Finance of Objectives,* 2018.