Università
Ca'Foscari
Venezia

Master's Degree programme — Second Cycle
(D.M. 270/2004)
in Economia e Finanza - Economics and Finance

Final Thesis

—

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Home Bias: PLS PM application on banking sector

Supervisor
Ch. Prof. Claudio Pizzi

Candidate
Enrico Piga
Matriculation number 856210

Academic Year
2016/2017

“*Life is always compared to a marathon, but I think it is more like being a sprinter, long stretches of hard work punctuated by brief moments in which we are given the opportunity to perform at our best*”

Michael Johnson

Abstract

This thesis aims to analyse banking sector in north-east of Italy, more precisely Co-operative Credit Banks connected with the phenomenon of Home Bias. In literature, many studies detected different causes of home bias both from a client and bank point of view, namely geographical factors related to widespread presence in small territories, cognitive errors in portfolio asset allocation caused by low diversification, political factors and costs in international financial markets. The thesis accompanies theoretical concepts with a quantitative analysis on a real dataset on which the application of Structural Equation Models, in the specific Partial Least Squared with different approaches. Furthermore, in light of heterogeneity in the dataset, we try to implement previous models in order to be in line with theory, obtaining homogeneous classes. Finally, mediating effect are considered and quantified as source deviation of relationships inside estimated model. The thesis is organised as follow. Chapter I introduces the basic concepts of Asset Allocation linked to home basis and ending with literature of possible causes. Chapter II describes banking sector and regulation in Italy. Introduction of PLS PM approach is exposed in Chapter III. Chapter IV contains the analysis of the dataset with descriptive statistics. In Chapter V PLS PM approach is applied on dataset. Heterogeneity in the dataset and the use of REBUS PLS to detect homogeneous classes are presented in Chapter VI. Chapter VII focuses on detecting possible moderating effects in PLS PM estimated on previous chapters. We conclude in Chapter VIII with comments on results and reflections.

Acknowledgements

Desidero ringraziare tutti coloro che hanno partecipato direttamente o indirettamente alla stesura di questa tesi con suggerimenti ed osservazioni: a loro va la mia gratitudine, anche se a me spetta la responsabilitá di ogni errore contenuto in questo lavoro.

Ringrazio anzitutto il Professor Pizzi per la disponibilitá nell'avermi affiancato durante la stesura della tesi e per i preziosi consigli tecnici. Proseguo con il Professor Rigoni, il quale ha messo a disposizione i dati che sono stati utilizzati in questo elaborato.

Desidero ringraziare i brillanti colleghi, compreso tutto il team di Venice Microfinance Initiative, con i quali ho condiviso gioie e momenti difficili della vita universitaria, con l'augurio di rimanere in contatto anche al di fuori delle aule universitarie. Ringrazio i miei coinquilini che hanno reso speciale l'anno accademico trascorso a Venezia e tutte le persone che hanno fatto parte di questi ultimi 2 anni universitari.

Vorrei infine ringraziare le persone a me piú care: gli amici con i quali sono cresciuto, sempre al mio fianco, conscio del fatto che che grazie a loro che sono arrivato fino a qui, Elena, citando Dalla " il motore della mia macchina", fonte inesauribile di affetto e quotidiano supporto per la quale una pagina di ringraziamenti no basterebbe ed in fine la mia famiglia, che mi ha permesso di frequentare l'universitá, mi ha incoraggiato a raggiungere questo traguardo e mai mi ha fatto mancare il suo amorevole sostegno.

Dedicato a tutti voi questo traguardo.

We made it!

Contents

Abstract ii

Acknowledgements iii

1 Home Bias 1

1.1 Asset Allocation Theory . . . 1

1.1.1 Capital Asset Pricing Model . . . 1

1.1.2 Diversification . . . 2

1.1.3 Home Bias as a specific case of diversification bias . . . 5

1.2 Home Bias Roots . . . 6

1.2.1 Behavioral Roots . . . 6

1.2.2 Political Roots . . . 8

1.2.3 Costs in international financial market . . . 9

1.2.4 Geographical incidence in Home Bias . . . 10

1.3 Evolution of Home Bias across years . . . 11

1.3.1 Opposite tendency: Euro Introduction and HB decline . . . 11

2 BCC in Italy and MIFID Restriction 12

2.1 CCBs in Italy and Economic influence . . . 12

2.2 Lending activity and Regulation for CCBs . . . 13

2.3 Cooperative Credit Banks: Features . . . 14

2.3.1 CCBs in Veneto: the "Federazione Veneta" case . . . 15

2.4 Future of CCBs in Italy . . . 15

3 Partial Least Square Path Model 17

3.1 SEM and PLS-PM . . . 17

Notation for Path Models 18
3.2 Outer model 19
3.2.1 Reflective indicators 20
3.2.2 Formative/ indicative indicators 22
3.3 Inner Model 23
3.4 PLS Algorithm 24
$1^{st} Phase$ 24
$2^{nd} Phase$ 25
$3^{rd} Phase$ 26
3.5 Model Validation 26
3.5.1 Measurement model validation 26
Reflective measures 26
Formative Measures 28
3.5.2 Structural Model Validation 28

4 Descriptive Statistics of the Dataset **30**
4.1 Dataset Analysis 30
4.1.1 Clients' infomations 32
4.2 Facilities Analysis 34
4.3 Performance Analysis 36
4.3.1 Return on Equity 36
4.3.2 Return on Assets 38
4.4 Qualitative Analysis 38
4.5 Home Bias 40
4.6 Data Cleaning 42
4.6.1 Proxy to use in computation 42

5 Hypothesis and methodology **44**
5.1 Brief literature of PLS-PM application in economics field 44
5.2 PLS PM Model Estimation 45
5.2.1 The model specification 45

5.2.2 Measurement model Assessment 48
5.2.3 Inner Model Assessment 51
5.2.4 Validation 53
5.2.5 Results' Assessment 54

6 Heterogeneity and REBUS PLS 55

6.1 Heterogeneity 55
6.1.1 The Rebus-PLS Algorithm 56
6.2 Home Bias segmentation and REBUS PLS application 57
6.3 Branches with High Home Bias 58
6.4 Branches with Low Home Bias 60
6.5 Analysis of Branches belonging to High or Low Home Bias 62
6.5.1 Elements of classes 64

7 Mediator and Moderator Effects 68

7.1 Difference between mediation and moderation 68
7.2 Moderated Regression Analysis 68
7.3 The Categorical Variable Approach 69
7.3.1 Geographical Proximity and Home Bias Relationship 70
7.3.2 Client's Risk Aversion and Home Bias Relationship 72
7.4 Mediator effect 72
7.4.1 High Home bias 73
7.4.2 Low Home bias 74

8 Conclusions 77

A Appendix 79

A.1 Tables of Chapter 4 79
A.2 Tables of Chapter 5 80
A.2.1 Bootstrap Validation Outputs 80
A.3 Tables of Chapter 6 81
A.3.1 Local Models - High Home Bias 81

A.3.2 Local Models - Low Home Bias 81
A.4 Tables of Chapter 7 82
A.4.1 Inner Model Assessments with Moderating Variables : Clients Risk Aversion - HB 82

B Appendix 83
B.1 Figures of Chapter 6 83
B.2 Figures of Chapter 7 85

Bibliography 86

List of Figures

1.1 Systematic Risk and Specific Risk 2
1.2 Investing pyramid 3

3.1 Notation used for parameters in PLS-PM 19
3.2 Path Model with Measurement model and Structural model 20
3.3 Loading and weights for reflective and formative measurement models 22

4.1 Veneto map and overview of bank's headquarters , 2007-2017 ©d-maps 31
4.2 CCBs' Number of clients for each bank and average number of branches from 2011 to 2015 35
4.3 CCB's Return on Equity from 2011 to 2015 37
4.4 Direct Fund Raising from 2011 to 2015 39
4.5 Asset Allocation 2014 client's portfolios 40
4.6 Graphic Home Bias form 2011 to 2015 41

5.1 Inner Model 49
5.2 Measurement model's loadings 50
5.3 Measurement model's weights 51

6.1 PLS PM - Inner Model of "High Home Bias" subset 59
6.2 Dendrogram of "High Home Bias" subset 60
6.3 PLS PM - Inner Model of "Low Home Bias" subset 62
6.4 Dendrogram of "Low Home Bias" subset 63
6.5 Graphical comparison between High Bond Home Bias branches and High Bond Home Bias branches 66

7.1 Moderators in the relationship between Geographical Proximity and Home Bias 70
7.2 Path Coefficient representation for mediation effect - High Home bias subset 73
7.3 Path Coefficient representation for mediation effect - Low Home bias subset 75

B.1 Local Model: Classes 1 and 3 83
B.2 Local Model: Classes 2 and 4 84
B.3 Local Model Low Home Bias: Classes 1 and 3 84
B.4 Local Model Low Home Bias: Class 3 85
B.5 Moderators in the relationship between Geographical Proximity and Home Bias 85

List of Tables

1.1 Home Bias roots literature 6

3.1 Parameter and Residuals notation for PLS-PM 23

4.1 Banking code and denominations of analysed banks 32

4.2 Clients' portfolio variables in the dataset 32

4.3 Banks' variables in the dataset 34

5.1 Check of blocks' unidimensionality 49

5.2 Home Bias Regression 51

5.3 Finaicial Performance Regression 52

5.4 Inner Model Assessemnt 52

5.5 Effects between latent variables in the Inner Model 53

6.1 Path coefficients and quality analysis: High Home Bias 61

6.2 Path coefficients and quality analysis: Low Home Bias 64

6.3 Path coefficients and quality analysis: Comparison High and Low Home Bias 65

7.1 Moderators effects - Geographical Proximity to Home Bias 71

7.2 Path coefficients for mediation effect on High Home bias subset . . 73

7.3 Validation test on indirect effect - High Home Bias subset 74

7.4 Path coefficients for mediation effect on Low Home bias subset . . . 74

7.5 Validation test on indirect effect - High Home Bias subset 75

A.1 Asset Allocation 2014 client's portfolios 79

A.2 Home bias across banks form 2011 to 2015 79

A.3 Weights Bootstrap Validation 80
A.4 Loadings Bootstrap Validation 80
A.5 Effects Bootstrap Validation 80
A.6 Path coefficients and quality analysis 81
A.7 Loadings comparison between global model and 4 local models . . 81
A.8 Home bias across banks form 2011 to 2015 82

List of Abbreviations

CAPM	**C**apital **A**sset **P**ricing **M**odel
CCB	**C**ooperative **C**redit **B**ank
SEM	**S**tructural **E**quation **M**odel
LV	**L**atent **V**ariable
MV	**M**anifest **V**ariable
ROE	**R**eturn **O**n **E**quity
ROA	**R**eturn **O**n **A**ssets
GQI	**G**roup **Q**uality **I**ndex

Chapter 1

Home Bias

1.1 Asset Allocation Theory

One of the most challenging problems economist have encountered across years is the measurement of the trade-off between risk and expected return. Literature produced dozens of statistical methodologies to assess possible investment strategies coherent with individual risk aversion.

Before introducing complex concepts about asset allocation and diversification, it worth to focus on the basis of the concept of equilibrium models. The purpose of a general equilibrium model is to "determine the relevant measure of risk for any asset and the relationship between expected return and risk for any asset when markets are in equilibrium"[1]. Thanks to equilibrium model we will be able to derive how portfolios should be constructed and give a reasonable idea of the optimal portfolio. It is crucial to bear in mind that models are a simplification of reality thanks to the adoption of many stringent assumptions and that best model is the one which fit better both market and individual behaviors.

1.1.1 Capital Asset Pricing Model

A turning point in portfolio theory was given by Markowitz, with the introduction for the first time of the Capital Assets Pricing Model (CAPM) which is a model that explain the relationships between systematic risk and expected returns.

The origins of CAPM are attributable to William Sharpe, Jhan Lintner e Jean Mossin, that from 1964 to 1966 putted the basis for one of the most important contribution in asset allocation methodology. The model is based on the idea that investors should be compensated for the time value of the money and risk: risk-free component representing the investment over a period of time and it is usually represented by the yield of a highly safe sovereign bond as US Treasury and German Bond. Risk part represents the compensation that investor should obtain to bear additional risk and it is calculated from the risk measure β, which compares asset risk with respect to market, and the difference between the risk of the most diversified portfolio on the market and the risk free.

1.1.2 Diversification

In economics, the concept of risk can be splitted in two components:

- Systematic Risk
- Specific Risk

The first reflects the movements of the assets due to changes in the market and the latter reflect the specific risk caused by factors or other variables addressable to the asset taken into consideration.

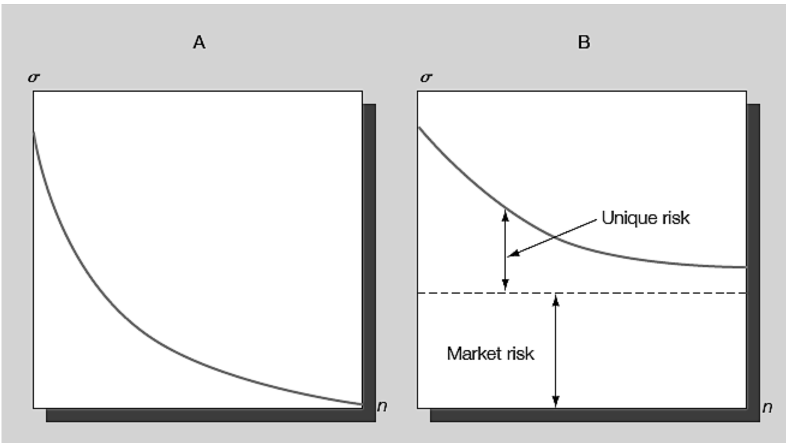

FIGURE 1.1: Systematic Risk and Specific Risk

The "systematic risk" cannot decrease under a certain level, whereas "specific risk" can be diminished thanks to an appropriate asset allocation and taking advantage of diversification. Following the so called "pyramid investing" exposed in Figure 1.2, the concept of diversification can be expressed from two different points of view: a horizontal diversification involves investing in the same asset (stock, bond, futures...) but referent areas; a vertical diversification involves investing in very different asset class from many different areas of competence.[1]

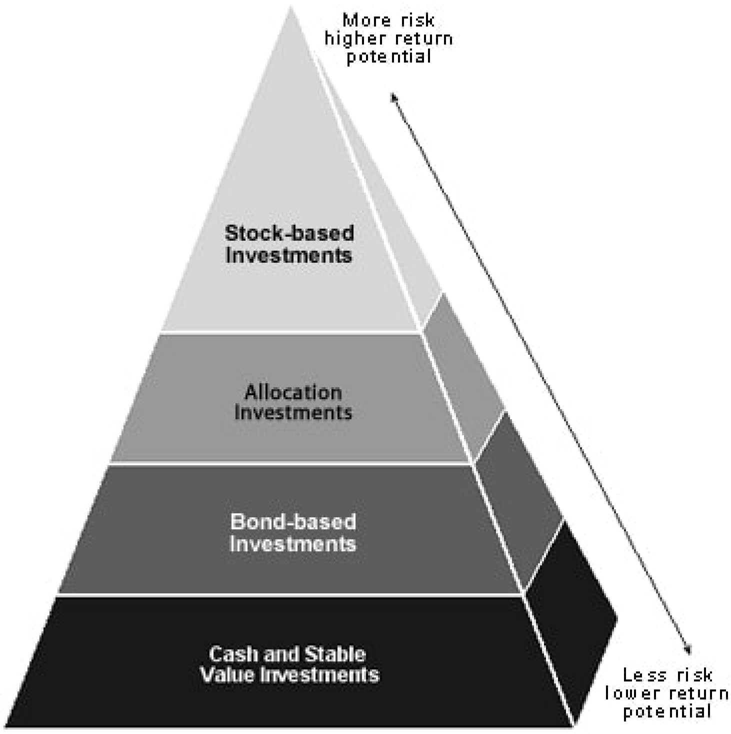

FIGURE 1.2: Investing pyramid

A similar subdivision of the risk has been made with a psychological approach from psychologist Lopez on 1987: a first "disposal factor" reflect the individual's nature behavior in front of a risky choice, whereas a second "environmental factor" describe limits and goals of the person. An asset is considered more desirable for portfolio diversification if systemic risk component just cover a small portion of the whole assets risk.

Asset allocation, driven by diversification principle, assigns more diversified portfolios to investors with a low risk profile and less diversified ones to those investors

[1]See also Portfolio diversification, Farlex Financial Dictionary. (2009). Retrieved June 23 2017 from http://financial dictionary.thefreedictionary.com/Portfolio+diversification

who present more attitude in bearing risk. In the next chapter, we will analyse personal attitude to bear risk and how international legislation fit the necessity to classify investors and their degree of financial competence. Investors can rely on two different ways to diversify their portfolio: adding titles with correlations lower than 1, add a higher number of titles in portfolio.

(a) Correlation is a measure of linear dependence between two variables, the output of that measure ranges from 1 to -1 respectively referred to strong positive correlation and strong negative correlation. A value equal to 0 displays no correlation at all between variables. Rational investors, considered a person who desires to maximize his/her needs, will be willing to insert in their portfolio titles with less than 1 correlation or negative to be less exposed to possible simultaneous falls.
Diversification assumes that the higher the correlation between titles the higher the risk. Thinking about a portfolio constituted only by stocks of one single segment of the market (i.e. metal market), in case of external shocks affecting the whole metal market the performance of the portfolio will be massively affected by a negative result without any doubts.

(b) From a theoretical point of view, assuming equal amounts are invested in each asset of the portfolio, with N assets the proportion in each asset is 1/N. As N gets larger, the variance of the portfolio gets increasingly smaller approaching zero.

Diversification is so important in asset management because it allows to reduce return volatility, maintaining the same expected return and "it is reasonable to pursue diversification as far as marginal costs are less than marginal benefits" [2].
Diversification bears also many drawdowns that make it expensive and extremely difficult to apply like costs involved in daily changes of portfolio composition.
It has also a behavioral "side" that reflects social-economical features of investors related to their personal characteristics and their risk aversion which can vary depending on the degree of risk. One of the main reasons of low degree of diversification lays on "Illusion of control" which is shown when an investor highly relies on

his abilities in selecting most performance titles. "An illusion of control could create an inappropriate level of over-confidence. Over-confident investors might choose not to diversify because they might mistakenly believe that they will earn superior performance by active trading" [3]. Authors verified this hypothesis affirming that investors affected by "Illusion of control" tend to change their portfolio composition maintaining a low number of titles, without benefit from diversification.

A more suitable reason of low degree of diversification can be express by prospect theory exposed for the first time by Tversky and Kahneman, which explain "the major violation of expected utility theory in choices between risky prospects with a small number of outcomes"[4]. A direct investment in stocks can lead to possible high returns, that may or may not happens with a relative probability, which cannot be possible in case of well diversified portfolio [2].

It is proved that investors are willing to bear risk in case of high return with low probability, accepting idiosyncratic risk. By contrast, there is risk aversion in case of low amount of losses with low probability. This can be interpreted as motivation for overweighting fixed-income investments.

A key points of this theory is that "a value function is concave for gains, convex for losses" and that "nonlinear transformation of the probability scale, which overweights small probabilities and underweights moderate and high probabilities" [5].

1.1.3 Home Bias as a specific case of diversification bias

Home bias, firstly studied by French and Poterba [6], is generally recognized as the "tendency for investors to favor or prefer companies based in their own country and overweight exposure to domestic investments in their portfolios" [7]. The Capital Asset Pricing Model (CAPM) predicts that, in frictionless financial markets, homogenous investors would hold a share of financial asset equal to the share of the financial assets of that country in the world portfolio [2].

A possible way to interpret home bias in international market is to measure individual allocation of domestic investments to their portfolio and compare it to the

[2] See also Cochrane (2005)

relevance of the invested country on the international stage. Theoretically, to obtain a well internationally diversified portfolio, allocations should be coherent with international country relevance.

Due to a low number of high quality data, the majority of research on home bias so far was conducted on equity and bond investments in developed countries.

In 2016, US equity marker size was 40% of international equity market size and, on average, American investors invested more than 75% [3] in the home country equity portfolio. It is evident how much investors are overexposed in relation to an internationally well diversified equity portfolio.

1.2 Home Bias Roots

Home Bias roots can be addressed to different reasons depending on individual behaviors, political pressure, costs on international financial market and geographical reasons linked to the concept of familiarity. There exist a wide literature about home bias roots and a brief summary is exposed in Table 1.1, where authors cited in this thesis have been divided by categories of roots.

TABLE 1.1: Home Bias roots literature

Roots	Authors
Behavioral	Kilka-Weber; Dziuda - Mondri
Political	Vanpée - De Moor ; De Marco - Macchiavelli
Int. Costs	Obstfeld - Rogof; Lewis; Domowitz - Madhaven ; Mann - Meade
Geographical	Ferreira - Miguel; Farquee; Boot

1.2.1 Behavioral Roots

Deviation from the optimal portfolio structure may arise from the idea that investors have more knowledge about the home market country instead of foreign markets. Kilka and Weber [8] conducted a survey of American and Dutch students about

[3] See also Coeurdacier, Nicolas; Rey, Hélene (2011/12/01). "Home Bias in Open Economy Financial Macroeconomics"

their expectation financial titles and indices's returns of their own country indicating a probability distribution of them, specifying how much they felt confident in answering.
Results highlight the tendency of the interviewees to provide more accurate predictions, in terms of probability distributions, as they felt more confident and they tend to over-estimate returns of the titles on which they believed to be more competent. These outcomes can be interpreted as investors prefer to give their opinion on subject they are familiar with to detriment of unknown topic. This can refer to optimism, in essence an overestimation of own abilities which leads predictions to be favorable to investors scope, and can be interpreted also as overconfidence, an underestimation of variability that in finance can be interpreted as underestimation of risks.
Thanks to the improvements of technologies and the increasing ease of access to financial markets for the individual investor, the problem of overconfidence is spreading. On-line trading can get access to financially unprepared and gambling individuals that are looking for "free lunch" opportunities exposing themselves to cognitive bias and overconfidence decision conditioning their trading performance. Information asymmetries can be a plausible cause of home bias for individual investors than for fund managers, who can devote a considerable amount of resources to gather relevant information. This idea should suggest that only the first category is affected by home bias.
It has not to be underestimated the fact that, only in the US, over 75% of the financial markets is controlled from mutual funds, pension funds and other financial institutions[4]. Dziuda and Mondri [9] wanted to highlight that even in professionally managed investment choices, there is presence of home bias. The authors distinguished individual investors, affected by asymmetric information, to professional portfolio managers, equally informed about all markets, and broadly considered able to generate "abnormal" returns. They divided into two periods the decisional activity: first period in which portfolio managers have to decide in which assets

[4]See also Commission on Corporate Governance of the New York Stock Exchange (July 23,2010, p. 12, http://www.nyse. com/pdfs/CCGReport.pdf)

investing between foreign or domestic assets, considering that they have to maintain the decision also for the next period. In the second period, individual investors conduct the asset-allocation based on managers' performances. Firstly, investors update their beliefs about managers based on first-period performance and due to their knowledge of the domestic market and they are willing to allocate resources to those managers that invest in domestic assets because they can better assess their performance. Secondly, managers are willing to intensify domestic investments to ensure custody of much more capital. As a result, in equilibrium the most skilled managers are willing to specialize in domestic assets which directly increase presence of home bias.

1.2.2 Political Roots

Around the end of the first decade of 2000, developed countries experienced a severe increasing of public spending and account deficits due to the well-known economic crisis. Form this scenario two different sources of home bias can arise:

- Supply-side drive
- Demand-side drive

The first one aims to contrast lack of funds, a massive issue of public debt as government bond was made. Domestic investors were the more natural subscribes and banks were obliged to hold government bond to be in line with capital requirements. The second one refers to market dynamics, leaded by higher returns guaranteed from an increase of interest rates as default state risk increased almost on every developed country [10].
Banks sovereign debt holders are considered the channel of transmission of credit risk from politically weak states to real economy, generate distortion in lending and investing activity. More precisely, in Europe, governments experienced tough times during sovereign bond crisis characterized by home bias problem. The phenomenon is evident from a massive presence of domestic sovereign bond in relative banks' portfolios, that can be valued as an exposure around 74%.

De Marco - Macchiavelli [11] wanted to get evidence of a different source of home bias, in practice, the political pressure exercised by former politicians sitting in the board of directors of some government-owned banks. The analysis was conducted mainly in GIIPS countries, namely Greece, Ireland, Italy, Portugal and Spain find a cross-sectional evidence of distortion with respect to both before and during the crisis periods.

A second analysis was conducted to highlight the moral suasion exercised by former politicians in relation with the amount of stock detained of the bank on which they were directors but they did not find statistical evidence.

Equity home bias has been widely studied across years, however the recent literature has intensified a new area of study on many different financial assets classes such as bond and corporate debt.

A relatively small amount of studies was conducted on bond home bias. One of the main reasons is that only in 2001, due to the contribution of IMF publishing bond holders' allocation across globe.

As a matter of fact, home bias on bond sector is even more pronounced than equity one. An interesting analysis was conducted on US bond market exposure, which is nearly 30% of the whole global market. US citizens destine on average over 90% for US fixed-income of their bond portfolio contrasting theoretical principle and exposing themselves to a relevant home bias problem [5].

1.2.3 Costs in international financial market

A relevant amount of literature review focused on relationship between asset allocation and costs in financial markets with different conclusions [12].

Obstfeld and Rogof [13] say that it is reasonable that asset allocation decision can be affected by transaction costs in real world increasing home bias relevance in financial markets. An opposite interpretation was given by Lewis (1999) affirming that home bias was not affected by transaction costs, related to a study on US investors.

[5] Source for data: Bloomberg. The World Federation of Exchanges, International Monetary Fund, data as of 31/08/2016

Domowitz, Glen, and Madhaven [14] discussed about the fact that transaction costs in US declined more than double, from 1996 to 1999, with respect to Europe. With a database of 42 countries, they use a panel data techniques and they interpreted the result affirming that investors tend to shift their asset allocation to countries which have low transaction costs.

1.2.4 Geographical incidence in Home Bias

Ferreira and Miguel [15] studied the behavior of bond home bias from 1997 to 2009 finding out that home bias is less present in "developed countries with higher restriction of foreign transaction, with higher developer bond market, higher familiarity and higher efficiency of judicial system".
With familiarity they mean, beyond other factors, the geographical proximity in the sense that investors felt more confident investing in a market in which they are more familiar with. Less informed investors tend to prefer domestic asset and "familiarity may be intended as cheap source of information".
Farquee [16] found a strong relationship in equity portfolio holders investing in countries that are geographically close. The concept of geographical proximity applied on investors with not enough financial background and lack of information can be highlighted on a national basis or even closer areas, such as regions. Another considerable factor that can influence in the asset allocation of small investors is the completely rely on local lending institutions, which very often are important players of economy in small realities.
These empirical evidences were called from Boot [17] "relationship lending" and it can be considered as one of the major points of strength of local institutions. Boot defines it as an activity "that:

- invests in obtaining customer-specific information, often proprietary in nature
- evaluates the profitability of these investments through multiple interactions with the same customer over time and/or across products"

The rest of the analysis about geographical factor will be analysed further with empirical case.

1.3 Evolution of Home Bias across years

1.3.1 Opposite tendency: Euro Introduction and HB decline

Schoenmaker and Bosch (2008) investigated on the effect of the introduction of the Euro as exchange risk can be considered one of the causes of Home Bias [18].
They found empirical evidence of a permanent decline of Home bias phenomenon on EMU countries whereas, as expected, no changes in no-EMU countries. The methodology used in this paper focused on the revelation of home bias 2 years before the introduction of the Euro currency and 2 years after that, in order to have an analysis more stable and trustworthy as possible.

Chapter 2

BCC in Italy and MIFID Restriction

2.1 CCBs in Italy and Economic influence

Banking activity in Italy is one of the oldest and important to the Italian economy. The birth of the modern banking sector can be temporarily placed in the Middle Ages, where the main activity was represented by deposit of goods and valuables. Even more importance to banks was given after the introduction of the concept of money, which started the creation of a secondary category of activities, namely lending function and monetary function. In Italy, banking sector developed much better than in other part of Europe, due to the presence of important commercial spaces as Venice and Florence and with other minor spots. The development of commercial activity and banking activity was increasing, leading Italy and its entrepreneurial network to a florid period in which the culture of banking spread around all Europe.

Also, modern stock exchanges find their origins in Italy, more precisely in Genoa and Venice, where bankers were highly present. For example, the Republic of Venice or the so called "La Serenissima", around 14th century was highly depending on money inflows to guarantee a stable growth to its commercial trade. Even though Italian operators were still partially devoted to the trade of products, they also turned into professional bankers and a remarkable development comes on both sectors.

A turning point in banking activity was given by Federico Guglielmo Raiffeisen, German politician, willing to help the poorest citizens of his region, asked to wealthier part of the region to lend them money, introducing the principle of unlimited responsibility of shareholders. Nowadays, the "Raiffeisen" model constituted one of the most important realities of mutualism and solidarity around the world. In Italy, following Raiffeisen model, around 1880 was born the first Cooperative Credit Bank (CCBs) near Padua, aiming to sustain socially and economically the growth of local communities. The mutual finality of these organisms aimd to ease entrepreneurs in their activities and one of the differences between non-cooperative institutes was that clients were also shareholders, aiming to increase cooperation and avoid independent and opportunistic behaviors. These entities found source of clients and shareholders in small communities, where there was a strict link with the territory increasing the trust of them.

2.2 Lending activity and Regulation for CCBs

As we said, lending activity was one of the oldest activities carried on by banks and it constitutes a high part of the total amount of liabilities from an accounting acceptation. Due to financial crisis of 2008, the banking sector was heavily damage form and economical point of view assessed with a billions of losses, and a loss of trust in banking system in general. Banks, in order to collect more capital, started issuing bonds, which were stimulated by a "decrease in taxes with respect to interest matured on deposit (12.5% for bonds and 27% for positive interests)"[19]. The lack of favorable inter-banking environment to collect capital forced institutes to have a more pressing from bank's financial advisors to clients to insert more CBB-issued bonds in their portfolios.

By the other hand, regulatory environment was obliged to discipline this practice in order to avoid illegal practice from advisors not in line with the ethical code of conduct. The MIFID (Market in Financial Instruments Directive), aimed to harmonize and guarantee higher transparency for laws around the EU's member states. MIFID I, was firstly enacted by the European Commission in 2004 with Directive

2004/39/EC[1], which aimed to the harmonization of the directive in all member states within 2008. With MIFID II, regulator aimed to increase transparency in banking activities and to increase the protection of investors, in consideration of the consequences of the financial crisis. In particular with this last goal, banks were forced to start profiling in a more accurate manner clients, in order to follow the principals of suitability and appropriateness of products suggested to clients.

2.3 Cooperative Credit Banks: Features

In 2016, the cooperative banks in Italy were over 330, which represented almost the 50% of the total amount of banking institutes and their activity represent over the 10% of the total loan market in Italy. Cooperative credit in Italy was subdivided on three different levels: local, regional and national where banks can associate in Federations of CCBs.

Cooperative banks were considered local banks known to be strictly linked to the territory and their clients. This form of localism let banks to increase the direct funding issuing more bonds to their clients and sustain trough fundings the development of financial and real economy. This could lead to the development of a form of bond home bias which was represented by the phenomenon of insert a high quantity of bank owner's account bond. The terms "high" was used to highlight the excess of presence of these instruments with respect to what asset allocation theory suggests. As the link between territory and banks, it was evident how much important was the role of the bank financial advisor which were the link between banks and clients.

Due to financial crisis of 2007-2008 a severe cut of lending activity stopped the whole economic chain and what was considered a financial crisis turned into a social crisis, in this phase the overall of the system suffered heavy losses. CCBs stable funds availability and liquidity gave them the opportunity to replace other banks, which

[1]Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on markets in financial instruments amending Council Directives 85/611/EEC and 93/6/EEC and Directive 2000/12/EC of the European Parliament and of the Council and repealing Council Directive 93/22/EEC

suffered much the credit crunch. More in detail, during 2009, in the explosion of the crisis,"loans to firms by the Italian CCBs have been grown by 4% while loans overall have decreased by 3%"[20], helping small realities to overcome that difficult moment.

2.3.1 CCBs in Veneto: the "Federazione Veneta" case

An important example of efficiency was represented by "Federazione Veneta", a network of CCBs spread in the northeast of Italy, which could rely on an entrepreneurial region known as one of the more productive in Italy with a high presence of small and medium enterprises. It was composed of 25 CCBs with over 126.00 shareholder which relied on the nature of cooperative credit as promoter and users of the services of the bank. It has a relationship with almost 700.000 clients assisted by a totality of 4000 employees[2].
As a matter of facts, if CCBs were able to lend money to a health entrepreneurial movement both citizens and banking sectors had to take benefit from that.
We will analyse more in deep the specific situation of eight of these banks through empirical study.

2.4 Future of CCBs in Italy

With the Law n. 49/2016 adopted in Italy, the aim was to reinforcement of cooperative banking system starting forcing mergers between CCBs in order to create cooperative banking groups, to increase the mutual goal and cohesion in the group but maintaining autonomy with respect to the correlated degree of risk. The goal is to create a new figure in the industry able to be competitive with non cooperative institutions, but holding high degree of presence on the territory. It could be considered an operational step to the European Economic and Monetary union as the EU is asking to its member states to increase banking size in order to obtain higher levels of efficiency and to cut costs tanking advantage of scale economies. Italy is

[2]Federazione Veneta - http://www2.fvbcc.it/Pagine/Home.aspx

following this directive and it started more competition on local and regional basis With the introduction of the new law, from an operative point of view, higher capital limit was established as maximum quote of participation in a CCB and the minimum number of shareholders to be considered a CCB increased from 250 to 500. The introduction of a "leader" among CCBs with a minimum capital of 1 billion and with a structure of a holding aimed to represent and manage each single CCB.
CCB had the possibility not to join the group, but guaranteeing some specific requirements about reserves and a minimum equity of 200 million.
There exist the general intention form the regulator to hold the mutual, territorial and cooperative nature of the CCB and contemporaneously develop a more safety and up to date structure, able to follow the development of the banking industry.

Chapter 3

Partial Least Square Path Model

3.1 SEM and PLS-PM

We can split statistics generation techniques in two time frames: around 80's multiple regression, logistic regression and exploratory factor analysis were developed mainly to deal with observed variables and with a general assumption that data were error free. A second statistic generation technique introduced around 1990, was referred to Simultaneous Equation Model (SEM) and can be considered much more than a statistical technique. It does integrate many different multivariate techniques into one model fitting framework coming from many different disciplines: measurement theory from psychology, path analysis form biology, factor analysis and regression from statistics and simultaneous equations from Econometrics. SEM is an advanced technique enables researchers to assess a complex model that has many relationships, performs confirmatory factor analysis and incorporates both unobserved and observed variables.
The main purpose of SEM is to provide and estimate parameters of the model, the variances and covariances of residual errors variances and parameters of observed variables. The second goal is to assess whether the model fits the data.
The Structured Equation Model has few approaches to estimate relationships:

- Covariance-based Methods (CB-SEM)
- Component-based Methods (PLS-PM)

CB-SEM has an estimation procedure based on maximum likelihood (ML) which "develops a theoretical covariance matrix based on a specified set of structural equations. The technique focuses on estimating a set of model parameters in such a way that the difference between the theoretical covariance matrix and the estimated covariance matrix is minimized"[21]. It also requires a quite stringent series of assumptions as the normality of data, a specific minimum sample size and this can be one of the causes of the strict number of researches using this methodology.
PLS-PM by its side, has an estimation procedure based on Ordinary Lest Squared (OLS) and aims to develop a model in which latent variables are as representative as possible of their own block of manifest variables. PLS-PM can be considered an iterative process which allows us to estimate weights, loadings and latent variable scores solved through multiple regression.
PLS-PM is considered to work efficiently with small sample size which involve complex relationships among variables. PLS-PM is preferred to CB-SEM method also for its greater statistic's power because it has the goal of minimizing error terms of endogenous constructs, maximizing the R^2 value, letting the model to have a very strong predictive feature.
It is a nonparametric method and provides robust model estimations, both with normal and non-normal distributed data. For all the reasons explained above, we are going to analyse at first theoretically and then empirically the use of PLS-PM method.

Notation for Path Models

PLS-PM is useful to specify systems of relationships rather than a dependent variable and a set of predictors; it is focuses on indirect (mediate) as well as direct effects of variable on other variables. The interesting feature of this model is the use of latent variables or construct variable, those which cannot be directly observed but can be measured thanks to manifest variables. Conventional notation for path model, as can be seen from Figure 3.1, uses oval boxes to indicate a latent variable, the item not directly observable; a rectangle box for observed variables, the manifest ones.

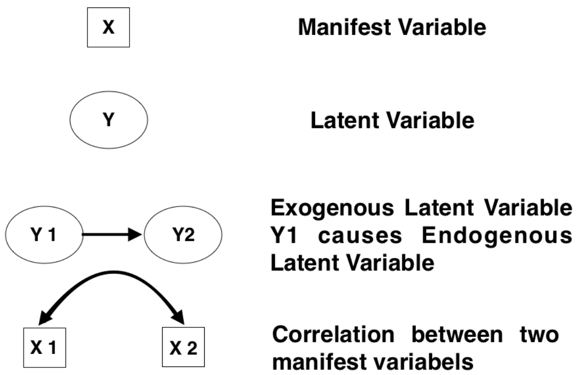

FIGURE 3.1: Notation used for parameters in PLS-PM

The error terms are linked to endogenous latent variable and to the manifest variables thanks to single-headed arrows. They represent the variance not computed from the estimated model. No error terms are present in exogenous latent variables. Error terms are present also in the structural/inner model, but are marked in a different way with respect to endogenous latent variable's error terms.
Benefits provided using latent variable can be the perfect suitable to a real-world scenario, indeed, as we can imagine, concepts in social sciences are complex and it can explain variation in human behavior.
PLS-PM can be interpreted as two distinct models, the structural model (inner model) and the measurement model (outer model) as is shown form Figure 3.2. The first is the model which just considers the relationship between only latent variables' class, the latter is the model which consider the relationship of each latent variable with their manifest variables and there can exist many in the model, usually recognized as blocks.

3.2 Outer model

Indicators or manifest variables can be divided in two categories, in relation with the causal relationships with the latent variable of the block:

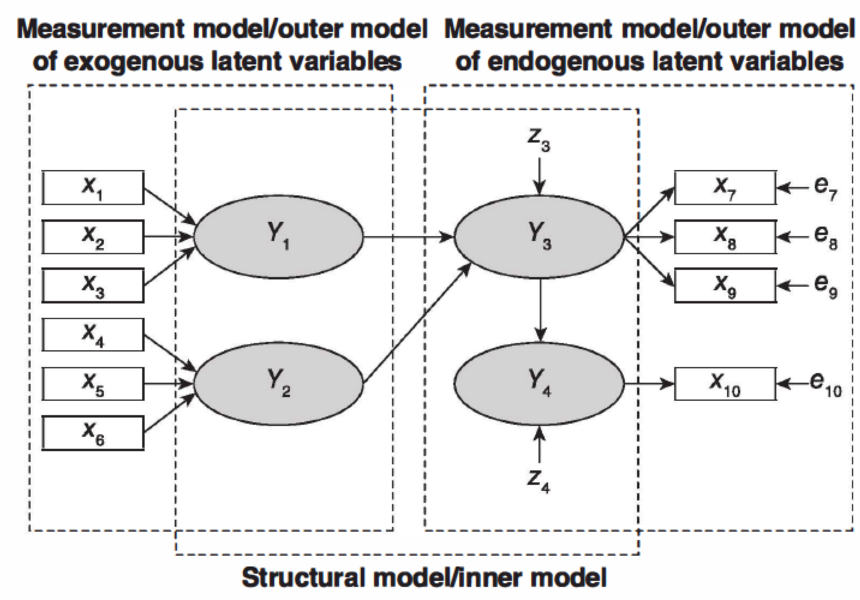

FIGURE 3.2: Path Model with Measurement model and Structural model

- reflective indicators: indicators which may affect latent variable
- formative indicators: indicators which can be interpreted as consequences of latent variable

3.2.1 Reflective indicators

Latent variable is considered as the cause of the manifest variables, indeed manifest variable is assumed to be a linear function of its latent variable:

$$x_{jk} = \lambda_{0jk} + \lambda_{jk} Y_j + \varepsilon_{jk}$$

where λ is called loading coefficient and ε_{jk} is the outer residuals of the regression. Is considered a regression through OLS, in this case we do not have any multicollinearity problem between indicators of the latent variable. In such a way, we can consider the latent variable as the main component of the block of manifest variables, under the idea which let manifest variables explain better the relationship with other latent variables.

This kind of indicators have specific requirements as any block must have unidimensional and homogeneous manifest variables.
In literature, in order to verify these properties were introduced 3 indices:

1. *Principal component analysis of a block*

   The block is considered unidimensional if the first eigenvalue of its correlation matrix is higher than 1

2. *Cronbach's alpha*

   Is considered a measure of internal consistency of a single block composed of p variables x_h (p is the number of manifest variables) when they are positively correlated each other. The block is considered homogeneous if this index is larger than 0.7.

   The formula can be expressed as follows:

$$\alpha = \frac{\sum_{h \neq h'} corr(x_h, x_{h'})}{p + \sum_{h \neq h'} corr(x_h, x_{h'})} \frac{p}{p-1}$$

3. *Dillon-Goldstein rho* Is also known as composite reliability and as Cronbach's alpha, the block is homogenous if index is larger than 0.7. Correlation sign is supposed to be positive between manifest variables x_h and its latent variable Y and this is obtainable if and only if λ loadings are positive.

$$\rho = \frac{(\sum_{h=1}^{p} \lambda_h)^2}{(\sum_{h=1}^{p} \lambda_h)^2 + \sum_{h=1}^{p}(1 - \lambda_h^2)}$$

If a variable violates these conditions, the block is considered non-unidimensional and feasible scenarios can arise to fix the problem. The removing of the manifest variable which do not let the block to be unidimensional, the splitting of the multidimensional block into smaller but higher explicative blocks or changing from reflective to formative block paying attention to the different interpretation of the latent variable.

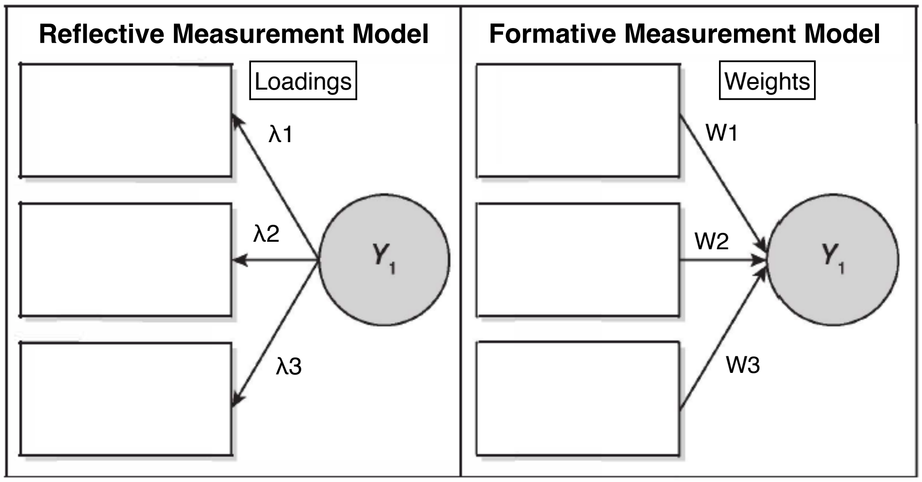

FIGURE 3.3: Loading and weights for reflective and formative measurement models

3.2.2 Formative/ indicative indicators

In this case, we are in front of a multiple regression OLS where some problems linked to multicollinearity[1] may arise. We can consider latent variable the best predictor of their manifest variable, under the assumption that lets manifest variables explain better the relationship between latent variable and other latent variables. Latent variable is considered as being caused by its manifest variables:

$$Y_j = w_{0j} + \sum_k w_{jk} x_{jk} + \delta_j$$

Weights are regression OLS coefficients obtained thanks to a multiple regression from a latent variable with their manifest variables. In case of multicollinearity between variables, a feasible way is to use loading instead of weights and using PLS regression coefficients.

[1]This phenomenon exist in multiple regression equations as far as an independent variable is considered highly correlated with other independent variables. The intrinsic problem arises because it can lead to misinterpretation of the statistical significance of an independent variable.

Parameters		**Residuals**	
Outer weights	w_{pq}	*Measurement model*	
Loadings	λ_{pq}	For formative indicators	δ_q
Path- coefficients	β_{mj}	For reflective indicators	ε_{pq}
		Structural Model	
		General	ζ_j

TABLE 3.1: Parameter and Residuals notation for PLS-PM

3.3 Inner Model

In the inner model, latent variables are related according to an underlying theory that want to be verified where latent variables can be interpreted as independent or dependent variables. In the first case we call them exogenous latent variable, when the variable is not caused by variables in the system, whereas in the second case we use the term endogenous when a latent variable depends on the behavior of other latent variables.

For each endogenous latent variable, measurement model can be written as:

$$Y_j = \sum_{h=1}^{p} \beta_{mj} Y_m + \zeta_j$$

where β_{mj} is the path coefficient which link the m-th latent variable to j-th endogenous latent variable. M is the number of latent variables which impact on generic endogenous latent variable, whereas ζ_j are the residuals of structural model. In Figure 3.3 is shown a sum up of the notation used for reflective and formative measurement models whereas in Table 3.1 parameters and residuals. This notation for Structural Equation Model will be used across the whole work in order to avoid misinterpretations.

3.4 PLS Algorithm

An important aspect of the PLS-PM method is the iterative algorithm used to estimate the model. A proper presentation and analysis of the different phases conducted from the algorithm is necessary to consolidate our theoretical background on PLS-PM.
The algorithm was first developed by Wold [22] and later on followed by deeper studies conducted by Lohmöller [23]. The original procedure can be subdivided into three phases as follows:

$\mathbf{1}^{st}Phase$

1. *Outer approximation of the latent variable scores*
   Outer proxies are calculated as linear combinations of their respective indicators and are standardised. The algorithm is iterative and the weight of the linear combination are the output of the previous iteration after the last step(step 4).

2. *Estimation of the inner weights*
   Inner weights are computed for each latent variable and give the idea of the strength of the relationships with other latent variables. In literature, three different schemes to the computation of inner weights have been presented, but usually researchers refer to the centroid scheme presented by Wold in 1982. Centroid scheme use correlation between the outer proxy computed in step 1 and adjacent latent variables, assign values of 1 or -1 in case of positive or negative correlation.
   The factor weighing is another scheme which take into consideration as weigh the value of the coefficients, highlighting the power of the relationship between variables.
   The more complete and exhaustive scheme is the path weight scheme, is referred to the relationship between latent variables and includes the advantage of considering the power of relationships and the causality between them.

The weights of those latent variables that explain the focal latent variable, is the correlation between them, whereas in case of latent variables explained by the focal latent variable, a multiple regression of all predictor variables of the latent one is run and weights are equal to regression coefficient.

3. *Inner approximation of the latent variable scores*
   Inner proxies of the latent variable are calculated as a linear combination of the outer proxies of near latent variables.

4. *Estimation of outer weights*
   The estimation differs from formative and reflective latent variable. In case of formative latent variables, weights come from OLS regression of the inner model proxy of each latent variable on its indicators. A different approach is reserved for reflective latent variables as weight are calculated as covariance between the inner proxy of each latent variable and its indicators.

The first phase is iterated until convergence is obtained, when inner model and outer model produce the same result of the latent variable. More precisely the iteration stops when a change in outer weights between two iterations is below a predefines stop criterion equal to $10^{-5}(i.e.0.00001)$.

2$^{nd}Phase$

When PLS algorithm converges, the final outer weights are used to elaborate the final scores for the latent variables as normalized weighted aggregates of the manifest variables.
Outer weights show the incidence of the indicator in the computation of the final scores for latent variable, a very high value for a weight means that the variable has a relevant incidence in the latent variable value. Outers loadings measure the correlation between indicators and latent variable.

$\mathbf{3}^{rd} Phase$

The last step is reserved to the computation of the path coefficients where the assigned scores are used to run a OLS regression to determine relationships in the structural models between variables.

3.5 Model Validation

Model validation is one of the crucial phases of the model utilization because with this procedure we want to check the accuracy and goodness though cross-validation of PLS-PM estimates and assess how well the theory fits the data. It is important to state that for PLS-PM does not exist a unique criterion to assess the goodness of fit, indeed different indices were built to evaluate the discrepancy between manifest or latent variable values and the values predicted by the model.
A specific set of indices is created to evaluate accuracy of measurement and structural model results using nonparametic evaluation criteria and thanks to techniques like bootstrapping [2]. The modus operandi of these criteria can be divided in two steps: assessing first reflective and formative measurement model and then, if the results are considered acceptable, the procedure follows with the assessment of the structural model results.

3.5.1 Measurement model validation

Reflective measures

These kinds of indicators help us to check if manifest variables are well explained by latent variables.

1. *Communality*
   It measures how much of the variability of the manifest variables is explained by latent variable of the block. With reference to block j the notation for communality index is :

[2]Bootstrapping is a nonparametric approach for estimating the accuracy of PLS parameters estimates through cross-validation

$$communality_j = \frac{1}{p_j}\sum_{h=1}^{p_j} corr^2(x_{jh}, y_j)$$

If we want to express communality average for all block the notation is:

$$\overline{communality} = \frac{1}{p}\sum_{j=1}^{J} p_j communality_j$$

that is:

$$\overline{communality} = \frac{1}{p}\sum_{h=1}^{p_j}\sum_{j=1}^{J} corr^2(x_{jh}, y_j)$$

where p is the number of manifest variables.

Thanks to this index, we check loadings which indicate the variance shared between the construct and its indicators.
If the indicator shows a low value of communality the variable under analysis should be excluded from the study.

2. *Average Variance Extracted (AVE)*

   This index is applied to explain the percentage of variance explained form each manifest variable with respect to the total amount of variance due to measurement model.

$$AVE = \frac{\sum \lambda_{jk}^2}{\sum \lambda_{jk}^2 + \sum var(\varepsilon_{jk})}$$

   The index should be larger than 0.5 which suggests that over 50% of the variance is accountable to manifest variables.

Formative Measures

Formative indicators behave in a different way with respect to reflective ones, indeed formative indicators do not necessarily measure the same underlying construct and for this they are not obliged to be correlated with each other. The indices used for reflective indicators cannot be used for formative indicators. We just should check which of the outer weight contribute most of the latent variable value. Caution should be made on formative indicators and the exclusion of some variables should be suggested in case of multicollinearity.

3.5.2 Structural Model Validation

For structural model validation researchers are used to rely on three measures:

1. R^2

   This value is computed for the endogenous latent variable, the interpretation is the same as in multiple regression analysis, indeed the value of R^2 aims to show the amount of variance in endogenous latent variable explained by its independent latent variables. The higher the value of R^2, the higher the goodness of the structural models.

2. *Redundancy*

   This index can be considered as a support of the R^2. it is not enough informative to know how much the regression can be adapted to the structural model, it is necessary to have some evaluation of the measurement model for each block of endogenous variables. Redundancy index measures the quality of structural model for each block of endogenous variables taking into consideration measurement model.

$$redundancy_j = communality_j * R^2_{j|Y}$$

where $R^2_{j|Y}$ is the R^2 coefficient from the regression between Y_j and its predictors Y_j. Redundancy average for all blocks of endogenous latent variables can measure the goodness of structural model:

$$\overline{redundancy_j} = \frac{1}{J}\sum_{j=1}^{J} redundancy_j$$

It measures the percent of variance of indicators in an endogenous block that is predicted from and independent latent variables associated with the endogenous latent variable. In essence, it shows how well a series of independent latent variables explain variations in the dependent latent variable. High values of redundancy can suggest a good ability of prediction.

3. *Goodness of fit (GoF)*

   Recently was introduced a so called "global index" for PLS-PM thanks to Amato (2004). It can be considered a unique index which can give the intuition of the goodness of fit of the model, considering both structural and measurement model. The index can be essentially considered a geometric mean of communality average and R^2 average. An acceptable value for GoF is >0.7 and we use $\overline{R^2}$ that is $\overline{R^2} = \frac{1}{J} R^2_{j|Y}$ in the notation:

$$GoF = \sqrt{\overline{communality} * \overline{R^2}}$$

Chapter 4

Descriptive Statistics of the Dataset

4.1 Dataset Analysis

The dataset we used for the empirical analysis was composed of CCBs and branches located in the northeast of Italy. The collection of these data was possible thanks to Prof. Rigoni and the Department of Economics of the University of Venice - Ca' Foscari. The dataset was composed by 8 CCBs and for each one a number of branches for a total amount of 166 bank statistic units and with 87725 clients from 2011 to 2015 that could physical and nonphysical individuals. For almost all statistic units, the dataset provided a range of 152 informative variables which could be divided into a few subsections in relation to their nature.
We mentioned only the informations which were considered useful for our empirical analysis making a deep descriptive study on manifest variables.
A first analysis of the whole dataset was conducted in order to have a complete overview of the set of banks and to carrying on a proper analysis I used in the PLS-PM model.
As we previously exposed on Chapter 2, banking sector in Italy was characterized by small bank groups which had a strong link with the territory and the people who lived in. The headquarters of these banks were geographically close each other and their presence was strictly referred to the Veneto's territory, as can appreciate from the map in Figure 4.1. Here we presented the analysis only for 8 of the 32 banks,

FIGURE 4.1: Veneto map and overview of bank's headquarters , 2007-2017 ©d-maps

which belonged to the "Federazione Veneta BCC[1]" that relied on 137.371 shareholders and 781.161 clients [32]. It was useful to have a geographical knowledge of banks' area of competence in order to understand better further analysis and considerations.

Above all, we were interested in analysing those banks which could guarantee us a good statistical significance in terms of quantity and quality. In Table 4.1 was shown a brief sum up of the banks under analysis of which we conducted informative variables reasonings.

[1]BBC stands for "Banca di Credito Cooperativo"

TABLE 4.1: Banking code and denominations of analysed banks

Baniking Code	Denomination
BACTD	BancAdria - Credito Cooperativo del Delta
BANSG	Banca San Giorgio Quinto Valle Agno
BMDSC	Banca di Monastier e del Sile Credito Cooperativo
BCDMR	Banca di Credito della Marca
BNMAR	Valpolicella Benaco Banca
CENMR	Centromarca Banca
CRAVC	Cassa Rurale ed Artigiana di Vestenanova Soc. Cooperativa
CREAR	Cassa Rurale ed Artigiana di Roana

4.1.1 Clients' infomations

A first set of data was referred to clients' wealth and instruments on which they invested in, giving an idea of the dynamics which might affect client's savings. The informative variables on Table 4.2, as all the others were collected with an annual frequency.

TABLE 4.2: Clients' portfolio variables in the dataset

Current Balance Account	Market Portfolio Value	Carrying Portfolio Value
CCB-issued Bank bond	Other-issued Bank bond	Government Bond Value
Stocks value	Funds/Sicav value	Insurance Policies Value
Derivatives Value	Repo Value	Certificate of Deposit Value

A brief explanation of the instruments which were present in portfolios was exposed below to understand the differences between these products and the underlying risk clients had to bear. It resulted particularly useful during model estimation where familiarity with products and related risk was necessary.

- Market Portfolio value : Was the equivalent value of the portfolio of each client at the moment of the valuation. It was calculated as the multiplication between the market value of each instrument and the number of relative instruments held by clients

- Carrying Portfolio value : Was the same concept of the Market Portfolio Value but instead of market value the value paid at the moment of the purchase was taken into consideration
- Current Balance account : Was the value detained by the client as liquid cash not invested in any particular instrument
- CCB-issued bonds : Equivalent value of the bond issued from CCB where client had the account[2]
- Other-issued Bank Bonds : Bonds issued from entities different from client's bank. It might referred to other bank bond, with exclusion of government bond
- Government Bond Value : Bonds issued from state government entities. Usually there existed many different types based on the maturity like US Bills, which matured in less than one year or U.S. Bonds that matured in more than ten years
- Stock values : Generic stock referred to many different entities. Could generate a profit for the owner in two different ways: capital gain and dividend payout. In case of positive difference between actual price of the stock and the one paid at the moment of the purchase we were dealing with a capital gain. Whereas, if company decides to distribute dividends to stockholders, the inflow of capital could be called dividend payout
- Funds/Sicav values : Dataset did not specify which type of fund are referring to and nothing about the riskiness of the instrument. In general, we the equivalent value of the total shares owned by the client and it was considered the number of shares owned multiplied by the value on the market of the specific fund

[2] A generic bond may generates profits for the client with coupon, paid by the issuer of the instrument (excluding the case of zero-coupon bond) and with the repayment of the borrowed capital.

- Insurance Policies Value : Was the value of insurance contracts owned by clients and signed with insurance companies
- Derivatives Value : The intrinsic purpose of this asset class was to hedge risks, its the value was linked to the value of an underlying assets. This type of investment should be suggested to those clients which are proficiency in financial markets, according to the MIFID survey, and are conscious of the risk they should bear once signed a derivative contract
- Repo Value : Was the value of the repurchase agreement investments. REPO was considered a form of short-term borrowing
- Certificate of Deposit Value: Nominative titles issued by a bank, which attribute the right to the holder to redeem the capital with interests. It was a type of investment that usually had a short-time horizon

4.2 Facilities Analysis

Informations inherent to the performance and health of the bank, summed up in Table 4.3 could be useful for further empirical analysis. We decided to split the analysis in 3 parts in order to highlight different aspects that will be used in PLS-PM model as manifest variables belonging to different blocks.

TABLE 4.3: Banks' variables in the dataset

N. of Employees as financial advisors	N. of Bank Offices
N. of Employees	N. of Bank's members
Total Assets	ROE
ROA	Direct Fund Rising

Banks' health could depend on many different factors, but in accordance with literature, the link a credit institute had with the territory was one of the most important ones. The more they are close to the clients the higher the trust in the institute should be.

A key factor and one of the most relevant influences in home bias phenomenon, according to the economics community, was the closeness to clients. It could be highlighted from the number of bank advisors on total number of employees, which putted in evidence the propensity of the bank to follow closer each client situation, providing a higher service in terms of quality. The underlying idea was that there existed a positive correlation between the number of clients and banks' geographical proximity to them clients, which can be measured by number of ATM or bank offices in each competence area. In addition, the propensity of the bank to provide a closer and complete service to clients might affect the decision to put savings in a bank rather than others.

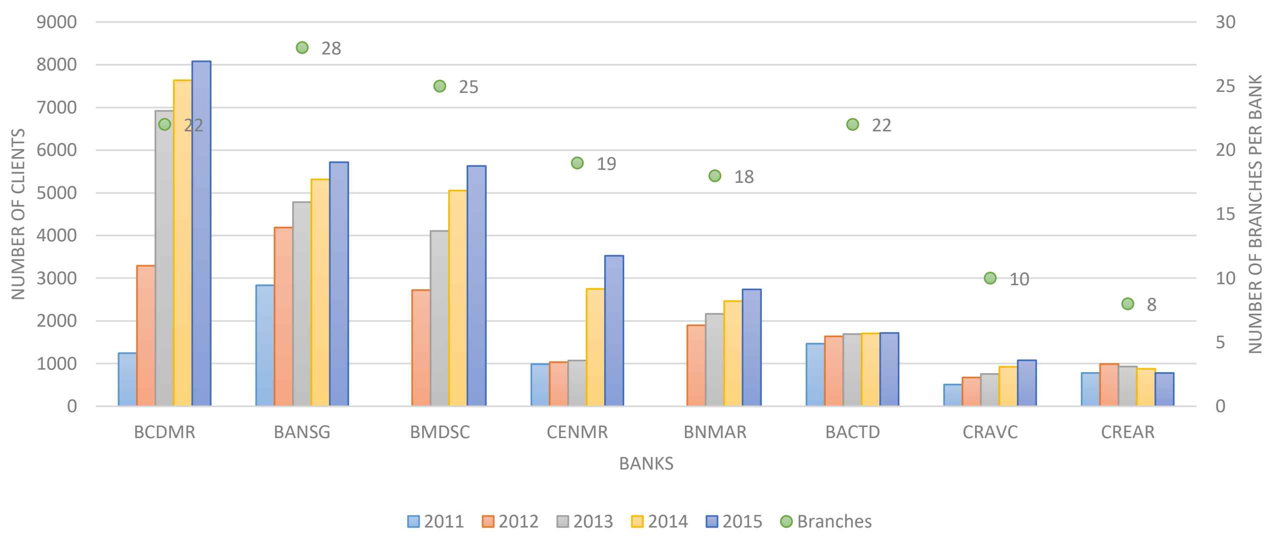

FIGURE 4.2: CCBs' Number of clients for each bank and average number of branches from 2011 to 2015

Taking in consideration Figure 4.2, clients increment was a constant almost for all banks, except CREAR bank which performed a decrease after 2012. If we assumed that the higher the number of clients, the higher the assets under administration, there could exist a linear dependence between number of clients and the presence on the territory of the bank represented by branches. All these factors, could be summed up as a geographical influence on the territory or geographical proximity to clients.

4.3 Performance Analysis

Banks' performances were directly summed up in the database as ROE and ROA. Now, we just considered these two indices even though we will generate a third factor which contributed in the next chapter in the creation of a single block of manifest variables.

4.3.1 Return on Equity

The return on equity was one of the most used financial indices which attests the ability of the banks to extract profits from its activity. Investors, those who believed in the management and their projects always look at high ROE which should convert in high dividends in future.
Generally, higher payout for investors means lower capital used to sustain the growth of the business. Supposing that home bias might increase when the percentage of own bank instrument, in our case CCB-issued bonds, with respect to the totality of fixed income instruments was higher than what economic theory suggests, profits and home bias mights be correlated. That should be a decision coming from bank management to increase own financing and the availability to use a wider amount of capital. The underlying idea was that high values of ROE might be linked with presence of home bias which may damage clients' accounts but had a positive influence for CCB.

$$ReturnOnEquity = \frac{NetIncome}{Shareholders'Equity} \tag{4.1}$$

A very important note had to be made in order to avoid misunderstandings: if there was a high presence of CCB-issued bond in clients' portfolios did not mean that there automatically was presence of home bias. If CCB's bonds was very advantageous under a risk-return profile, considered better than bond presented on the market, clients could invest all her/his money on that but just because it should be coherent with asset allocation theory.

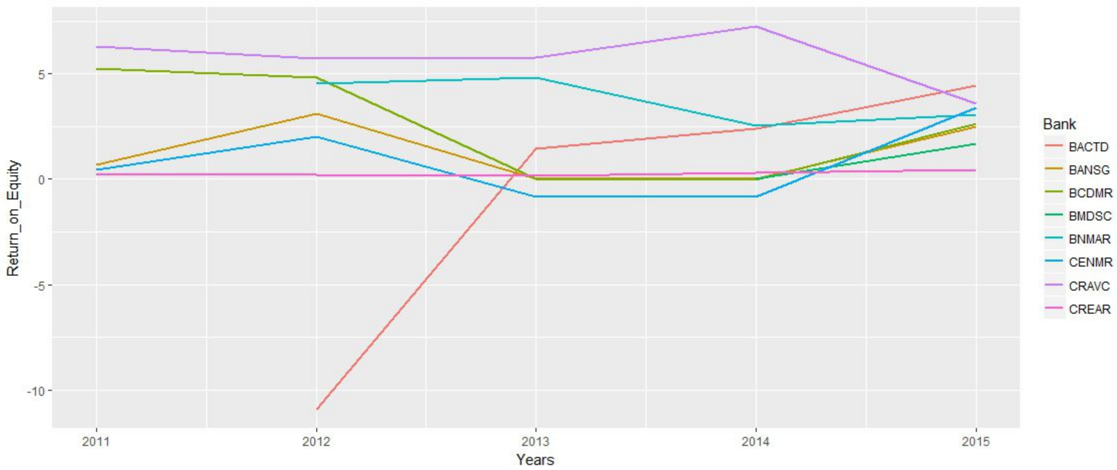

FIGURE 4.3: CCB's Return on Equity from 2011 to 2015

In order to conduct our empirical analysis, we made the assumption that CCB-issed bonds and other bonds on the market have the same risk and from a client point of view there should not be any differences in selecting one instead of another.

A brief exposition of the ROE's performances from 2011 to 2015 is shown in Figure 4.3.

As we can see, the overall level of ROE was constantly around values between 5% an -5%. It was a usual result for mutual credit institute, due to their structural nature of small institutions and clients with a low risk profile.

A separated analysis for BACTD ROE performance should be necessary to understand the 2012 results. It seemed that at the beginning of the data collection a very negative ROE could be the effects of some missing data, in a different case we could affirm that BACTD was in a severe financial distress situation since 2012 and maybe, thanks to a good investment plan and a slight increase of the number of clients, from 2012 onward the bank performed 3 consecutive positive ROEs, making the best profitability increment across all banks under analysis. CRAVC registered the best performance in 2014 with a ROE higher than 6% whereas all the other competitors performed below 3%.

4.3.2 Return on Assets

The return on asset was another important index on which managers relied on. Is the ratio between net income and total assets as can be seen from the formula below.

$$ReturnOnAssets = \frac{NetIncome}{TotalAssets} \tag{4.2}$$

An interesting topic about ROE and ROA could arise when there was a misalignment. From a mathematical point of view, formulas of ROE and ROA were the same if we assumed that a company did not use debt as a source of financing. Starting from this situation, we could imagine that if the company started leveraging, without any reference of which type of debt was considered, it would use the availability of capital to increase the assets, which followed an increase of asset value from an accounting point of view. The denominator of ROA was heavier and, assuming that net income remained constant, the index started decreasing. From the other hand, equity value was smaller, leading ROE value to increase. As a matter of facts, there was a negative correlation between indices and the increase of CCB-issued bond might be suitable for this situation and might let us see a possible influence of ROE/ROA behavior of the home bias phenomenon from a banking point of view on an empirical basis.

4.4 Qualitative Analysis

As we know from Chapter 2, fund-rinsing activity played an important role in banking activity. To understand the bank policy regarding fund-raising, resulted informative to check the evolution of direct fund-rising and indirect fund-rising. An increase in direct fund-raising, which could be seen from Figure 4.4, was the result of issuing CCB's bond. An example of such behavior could be represented by BCDMR performances; direct fund-rising increased from 2011 to 2014 and performed a slight

decrease in 2015 with respect to the previous year.
During 2014, as highlighted by clients' portfolio composition in Figure 4.5, the percentage of the CCB-isseud bond and the presence in CCB clients' portfolio was around 45% [3].

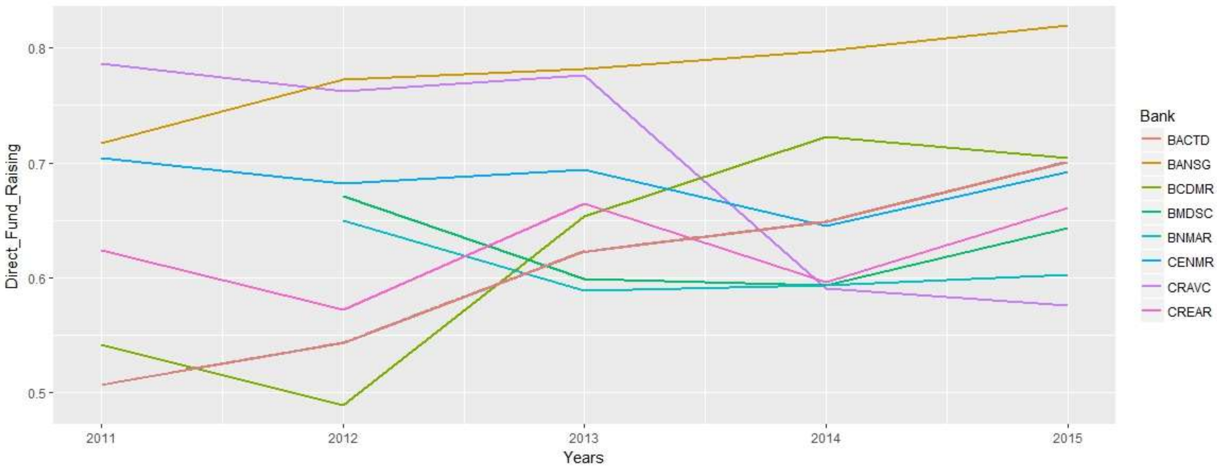

FIGURE 4.4: Direct Fund Raising from 2011 to 2015

In Table A.1, available on Appendix A, was shown a proxy of portfolios assets allocation for the medium client in each bank. For convenience, we took in analysis only asset allocation in 2014 considered that in other years many branches presented lacks of data. In general, we could affirm that the instrument much more owned by clients was the CCB-issued bond (bcc.bond) which had an overall incidence around 40%, except for BACTD on which the highest class concentration was represented by government bond (Gov.bond). High risk class seemed not to be favorable from clients because in means both stocks and derivatives did not exceed the 6% incidence with a very marked aversion for high risk as derivatives were never higher than 1%.
As we could imagine, the type of clients who had relationships with cooperative credit banks were not aiming to obtain high returns from their investments rather

[3]The percentage values are the result of the mean of all the portfolio present in each branch

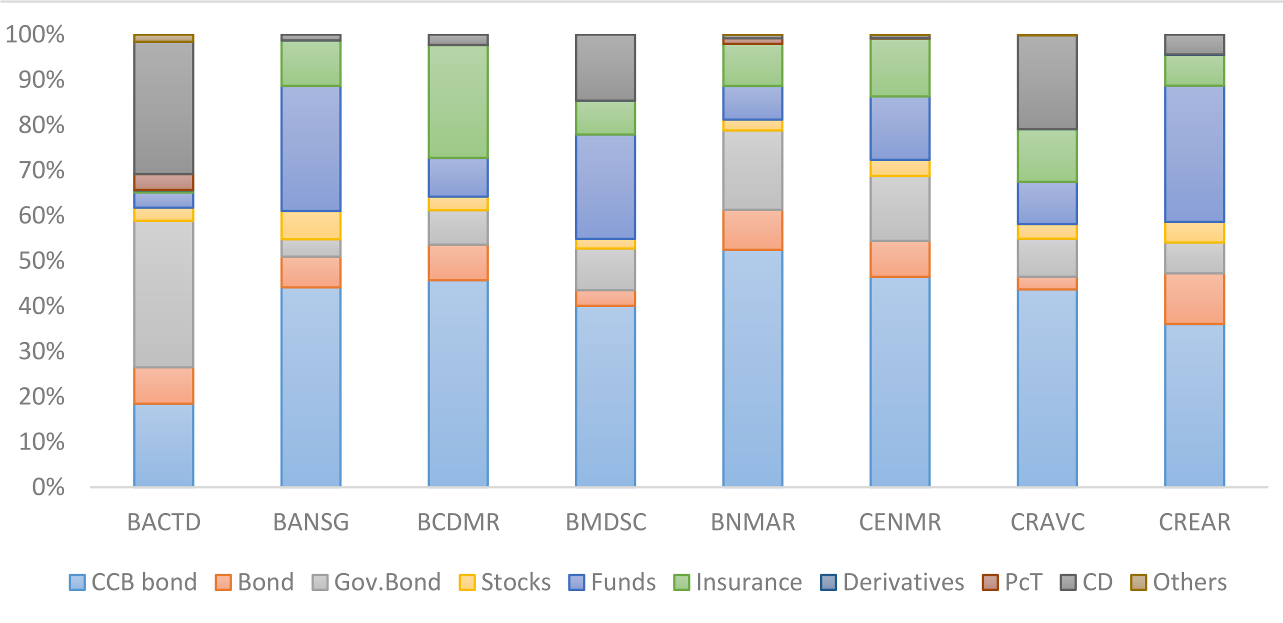

FIGURE 4.5: Asset Allocation 2014 client's portfolios

they wanted to hold an adverse risk profile choosing low risk investments. This idea was consistent with the portfolio composition as low risk investments represent a high percentage.

4.5 Home Bias

Home bias, as we were used to know, was considered a cognitive bias which led to low performances of client portfolio. Now, thanks to a change in the point of view, this phenomenon obtained a different meaning. Using clients as statistic unit, presence of home bias in their portfolios was something to avoid, but if we used branches of a bank as statistic unit, an increase of home bias could directly affect the availability of money for the bank. The issuing of new debt for a bank could directly affect financial performances and taking advantage of leverage could influence asset and liability composition.

A first indicative indicator was represented by Home Bias index, it composed as:

$$HomeBias = \frac{bcc.bond}{bcc.bond + bond + gov.bond + cd} \tag{4.3}$$

This indicator ranged from 0 to 1 and the maximum value indicated that in means, clients of a single branch were completely exposed to home bias problem. In case we were referring to a bank, the value represented how much an average client was affect from home bias.

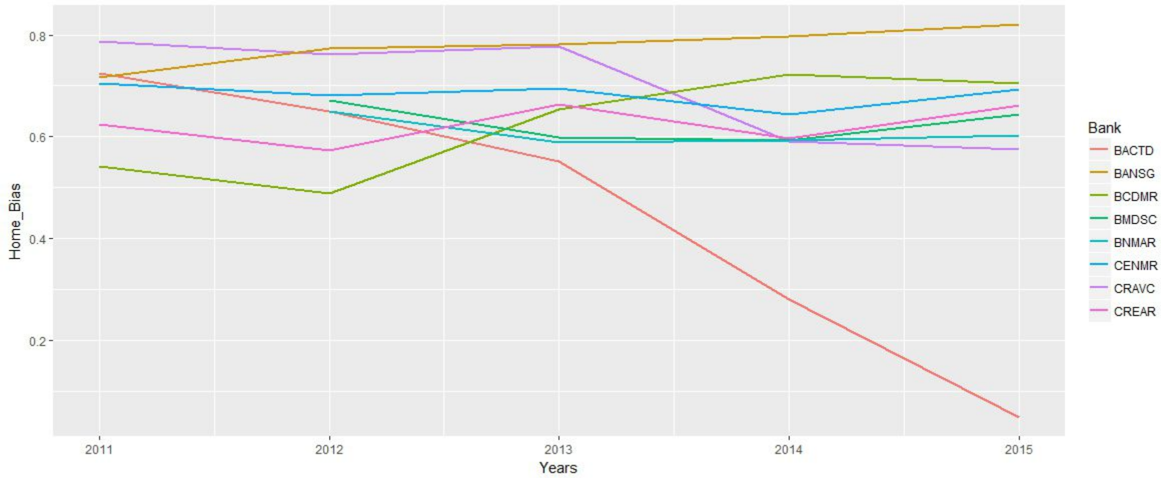

FIGURE 4.6: Graphic Home Bias form 2011 to 2015

As presented in Figure 4.6, trends of home bias index for all the banks in consideration had different behaviors. There was presence of home bias in all the analysed banks but the trend across years was different between banks, shown on Table A.2 available on Appendix A. CREAR, which presented a very high presence CCB-issued bond in their client's portfolio, showed an HB from 2011 to 2015 over 0.75, which verified the link. By contrast, other banks had different behaviors, like CRAVC which had a drastic change in 2014 where it performed a very slow value of the index.

It was interesting to understand if high values of home bias, associated to a theoretical disadvantage for clients due to a deviation from canonical asset allocation theory, might affect banks' performances. With this idea, we can highlight those banks which conducted a fair and trustable advisory activity with respect to those ones which may inject own bond in a client's portfolio to obtain better performances.

4.6 Data Cleaning

As far as our analysis used branches as statistics unit, we deliberately decided, facilitating the analysis process, to use the sum of balance account value, market account value and carrying portfolio value.
The relative MIFID documentation filed by each client were not useful for our analysis as too many observations on a restrict range of selection would be misleading and of difficult understanding. A different approach was used for the instruments present in the clients' portfolios as we took the mean of each of them for each branch of the bank. The informations had been taken with the same time frame and it was indicative of the tendency of branches to push to their clients product of the delegated bank. Essentially, this data could help us in understanding better the intrinsic causes of home bias in the banking sector and its influence in bank's financial performances.

4.6.1 Proxy to use in computation

We decided to create few variables that will result useful in empirical analysis. Here we just wanted to give an explanation of them and motivate such use.
Our interest was focused in estimate a proxy of the total assets of each branch of the bank. That would allow us to create new variables that might contribute our analysis. To do that, we based our process on the hypothesis that direct raising fund made by a bank can be estimated in a range between 75% and 85% of total assets of the bank. So, we used this information to estimate the total assets for all branches of our interest as proxy of the real total assets with an 80% incidence. Thanks to this passage, we were able to asses ROA and ROE index for each branch that would contribute as manifest variables in the estimation of latent variable "Performance". Another very important index used to assess bank performance is ROE (Return On Equity). We only had CCB's ROE but we want to highlight which branch contributed more in ROE proxy value. It would help us in finding which branch can be considered as "influencer" in home bias phenomenon and then making a more

accurate and deeper analysis.

The following Formula 4.4, was created to give an idea of the ROE, profitability measure, linked with the dimension of the branch and with the total amount of carrying portfolio value; in addition we inserted a penalty for a higher number of clients. In that way, we customized a single ROE for each branch attributing much more importance (higher ROE value) to those branches which were able to have high value of portfolio, but with a relatively low number of client, those branches which deserved higher ROE and which contributed more to the overall ROE.

$$ROEproxy = \frac{[(\frac{BankROE}{PorfolioBankMktValue})PortfolioBranch]/\frac{Branch.n.clients}{TotalBankClients}}{[\frac{BankROE}{PorfolioBankMktValue}PortfolioBranch]BankROE} \tag{4.4}$$

Other indices created ad hoc will be illustrated later on, during the model presentation as many of them made sense on model specification, due to possible different interpretation we gave to the latent variables. The composition of the latent variable and their links will be explained in the next chapter with a direct application to the dataset.

Chapter 5

Hypothesis and methodology

5.1 Brief literature of PLS-PM application in economics field

The main goal of our study was to measure home bias phenomenon, related to bond asset class, across the banks analysed so far and to understand if the cause-effect relationship we assumed on a theoretical basis could be confirmed or denied from the results.
An important contribution of the PLS PM in economics field was the research made by Trinchera and Russolillo [30] which found the necessity to describe even more "complex phenomena like poverty, progress, well-being…summarizing complex and multidimensional issues". Their aim was to provide composite indicators used in a modified version of PLS algorithm to describe, with the use of latent variables, the economic reality in a better way.
One of the first application of PLS PM in economics and business was given possible thanks to the contribution of Serrano-Cinca, Fuertez-Calleń, Gutieŕrez-Nieto and CueĺIar-Fernandez [31] in 2011, which analysed accounting data to assess causes of bankruptcy of US banks since 2009.
Furthermore, thanks to the contribution made by Danilo Frare on his Master Thesis "Profilo degli investitori e asset allocation degli investimenti finanziari ", using the same dataset analysed in this study, was possible to draw important conclusions about asset allocation. Individual characteristics with the results of the MIFID

survey for each bank clients led to outline common factors which might affect the composition of the portfolio. The these results were possible thanks to PLS PM model having individuals as statistical units. Main conclusions were represented from the existence of a positive correlation of financial knowledge and experience with investments in stock markets, the higher wage of the client seemed related to lower percentage of the portfolio allocated in risky assets, assessing wealth investors as "risk lover". A border effect on asset allocation is also made by holding period and age of the clients.

In addiction, it was important to highlight the study made by Giorgia Simon, who discovered the main factors which led households investment in bank bonds [25]. Three main causes drove investment decisions, namely familiarity investments, overconfidence and branch-level advisor characteristics. The empirical results were obtained thanks to the use of Fractional Response Model (FRM) with robust standard error and it was made using the same dataset we used in this thesis. We tried to increment the validity of the obtained results and to contribute as possible to amplify the literature of the bond allocation in Italian CCBs.

The aim of this chapter was to understand the home bias phenomenon and policies decisions taken in CCB's headquarters using the PLS-PM approach, where banks' branches were considered as a statistical unit.

5.2 PLS PM Model Estimation

The regression analysis of the dataset was based on the Partial Least Square - Path Model thanks to the support of the package plspm() available on the statistical software "R". The study of the model is articulated in 4 steps: the model specification, the measurement model assessment, inner model assessment and validation.

5.2.1 The model specification

The first step was considered one of the most important as the relationships that existed between groups of manifest variables and latent variables might affect the

results and fitness of the model.
We introduced in the model 2 exogenous latent variables, namely "Clients' Risk Aversion" and "Geographical Proximity", and 2 endogenous latent variables namely "Home Bias" and "Performances".
All the considerations made about manifest variables were referred to an average bank, meaning that the institute is managed in a fair way respecting MIFID capital restrictions and no situation of financial distress was affecting the bank at the moment of the data collection.

Exogenous Variables We considered in our model "Clients' Risk Aversion" as a reflective latent variable, knowing that the latent variable was considered the cause of the associated manifest variables. Manifest variable taken into consideration were represented by:

- "Per.bcc": it was the percentage of CCB-issued bonds with respect of all the other asset classes present in portfolio clients. It was a proxy of the relevance of this type of assets.
  High percentage value would be interpreted as that portfolio is not well diversified as too much capital is concentrated in a unique asset class with important concentration on safety assets
- "Per.low.risk": it was the percentage of low-risk assets[1] in portfolio clients.
  It was a proxy of the riskiness hold by an average client. High percentage value suggests that on average clients felt confident in no risky assets

The second exogenous variable of the model was represented by "Geographical Proximity" as reflective latent variable. Its manifest variables were represented by:

- "n.risorse" : number of employees without the specification of the role. The variable was considered for each branch and the higher the number of employees, the higher should be the attention for the client. It could also be

[1] CCB-issued bond, other bond, government bond and CD

broadly considered a measure of a bank's health as a high number of employees meant high costs for the bank. Assumed that, in the banking sector, usually when an institute suffers financial problems aims to reduce the staff, if the employees number remains constant, it may be interpreted as a proxy of good financial condition

- "n.inv.emp": it was the number of branch financial advisors. A high value meant that in the specific branch there was the availability of resources to follow a high number of clients and with more efficiency. We might also think that branches with high value of financial advisors had also more capability to encourage clients to buy CCB-issued bond
- "branch.client": it was the total number of clients for each branch

Endogenous variables We created two endogenous latent variables, both affected at least from 1 endogenous and 1 exogenous latent variable.
The first endogenous latent variable was "Bond Home Bias", which represented the presence of a home bias phenomenon in each branch for the bond asset class. Is a reflective block, due to the nature of endogenous latent variables which cannot be represented in a formative way and it was represented form 3 manifest variables namely:

- "HB": was a proxy of the home bias calculated for each branch as the ratio between CCB-issued bonds and the sum of all low risky assets present in the portfolio, as shown in Equation 4.3
- "bcc.on.mkt.val": was the incidence of CCB-issued bond on portfolio market value which wanted to highlight the incidence of this specific asset
- "bcc1": percentage of CCB-issued bonds on total bond asset class[2]. It showed how much the client was exposed to bond with a single institute and how much the bank could obtain fundings from each client, in relation to the client's willing to invest in bonds

[2]total bond constituted by CCB-issued bonds and other bonds

We believed that Home bias was not only affected by ratio in Equation 4.3, but might be directly influenced from other asset allocation dynamics and this was the reason why took more than variable HB into consideration.
The second endogenous latent variable was "Financial Performances", an indicator of bank's health and economic performances affected form a reflective block of 3 variables represented by:

- "ROA": Return on Assets of each branch
- "ROE": Return on Equity of each branch
- "Netincome": Net Income of each branch

We decided to use ROA and ROE, which had very high positive correlation (0.9442), even though from a theoretical point of view, as we said in Chapter 4, the two variables should have an opposite behavior, showing a negative correlation in case of CCB-issued bond increments in portfolio clients. We could interpret this result as possible Assets & Liabilities dynamics not captured from the data and not shown from the ratio empirical results.
In figure 5.1, we could see the inner model obtained after run the PLS PM algorithm, which putted in evidence only latent variables and their relationships. We thought a direct influence of "Clients' Risk Aversion" and "Geographical Proximity" on "HB" with another double direct effect of "Geographical Proximity" and "HB" on "Financial Performances". The exogenous variable were the main factors, form a literature point of view, which affected home bias phenomenon and our main goal was to discover the impact from these variable to CCBs' performances.

5.2.2 Measurement model Assessment

The assessment of each block was divided into 3 different phases: first it was checked the unidemsionality of the indicators, than it was checked if the indicators were well explained by its latent variables and finally it was assessed the degree of uniqueness of a given block from others.

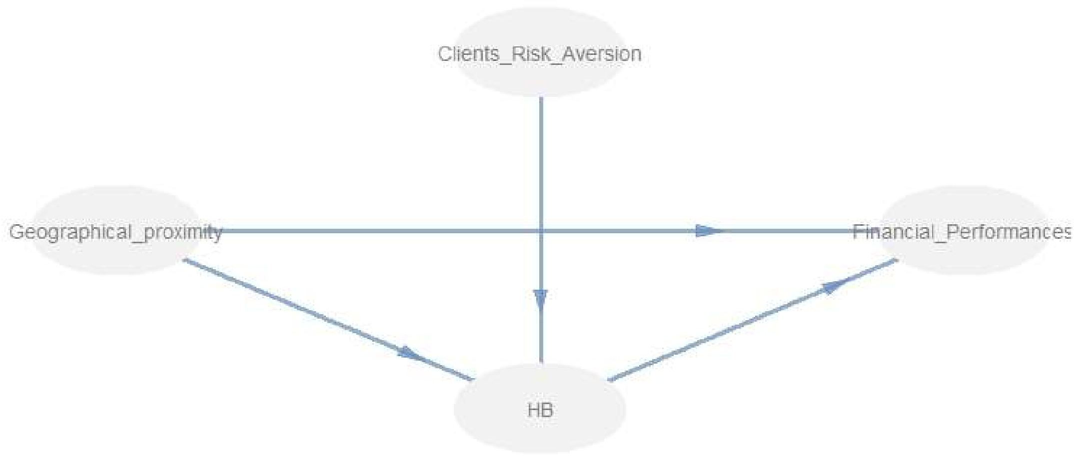

FIGURE 5.1: Inner Model

The unidimensionality of each reflective block was assessed thanks to the analysis on Table 5.1:

TABLE 5.1: Check of blocks' unidimensionality

	Mode	MVs	C.alpha	DG.rho	eig.1st	eig.2nd
Clients' Risk Aversion	A	2	0.9239584	0.9633718	1.858664	0.1413357
Geographical Proximity	A	3	0.8466390	0.9077234	2.300029	0.4809430
HB	A	3	0.8870994	0.9303453	2.450415	0.3861394
Financial Performances	A	3	0.9037503	0.9406414	2.524053	0.4209110

Cronbach's alpha and Dillon-Goldstein' rho were both higher that the acceptable level of 0.7. First and second eigenvalue were respectively higher and lower that level 1 and in general performed very distant values. In other words, we could say that all blocks in the model were considered unidimensional and that manifest variable in each block were positively correlated each other.
Measurement model assessment procedure took into consideration loadings and weights of the model, which helped us in understanding the goodness of the blocks.

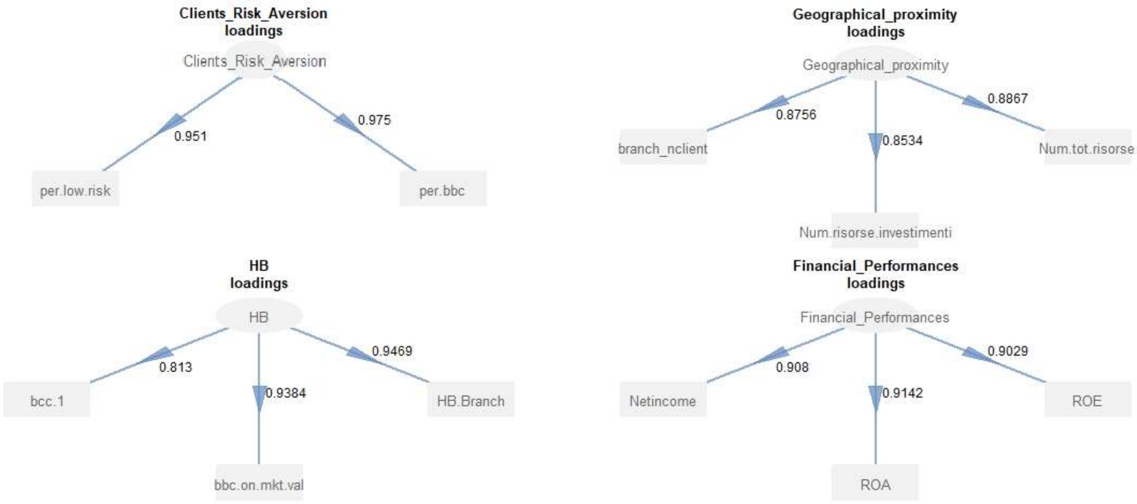

FIGURE 5.2: Measurement model's loadings

As we know, loadings in Figure 5.2 measured the correlation between indicators and latent variable and in our case we had a quite relevant correlation between variables.

Weights inside blocks, exposed in Figure 5.3, were well distributes as no one variable has an incidence considered "dangerous" in latent variable score. In case of a very high value of the weights, we might think that others variables were bringing low significance to the latent variable and might be useless to insert in the model. Communality of all variables, referred to their own blocks, were high enough to say that they were informative for the model and we should hold them to follow the analysis.

As last step of measurement model assessment, we needed to check cross-loadings between different blocks. The reason why we did this was to check if there were any variables which might have higher loading value related to a different block. This should make us think about the composition of the block and its specification as one or more indicator were misrepresenting the latent variable.

In conclusion, we assessed the measurement model appropriate and in line with acceptable results.

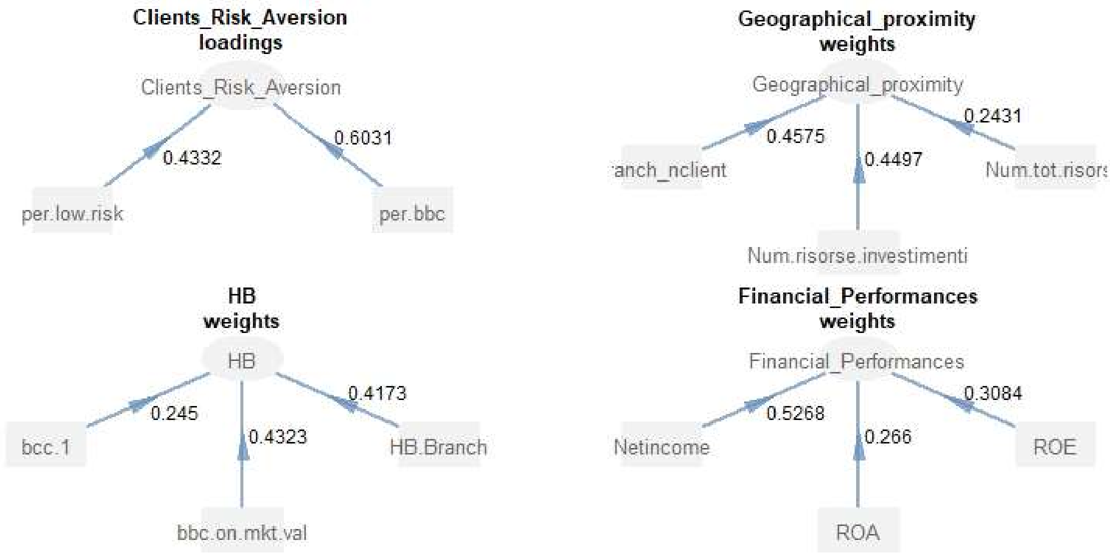

FIGURE 5.3: Measurement model's weights

5.2.3 Inner Model Assessment

The inner model assessment started putting in evidence the endogenous variables of the model and making a linear model, one for each endogenous variable, with exogenous variables to understand the sign of the β[3].

TABLE 5.2: Home Bias Regression

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-1.595128e-17	0.04714104	-3.383735e-16	1.000000e+00
Clients' Risk Aversion	7.918981e-01	0.04773658	1.658892e+01	7.803218e-37
Geographical Proximity	3.721486e-02	0.04773658	7.795880e-01	4.367639e-01

In the first equation, presented in Table 5.2, betas estimated for the exogenous variables have positive sign meaning that there was a positive impact on HB, the endogenous variable. Together with coefficient estimation, Std.Error and T statistics were exposed and should help us to see if coefficients were considered statistically significant, giving some intuition about the effects on the dependent variables.

[3] β on linear regressions express whether an exogenous variable had a positive or negative relationship with the endogenous variable and its incidence on the outcome variable in case of statistical significance

TABLE 5.3: Finaicial Performance Regression

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.076222e-17	0.073437	1.235912e-15	1.000000e+00
Geographical Proximity	-2.986163e-01	0.074419	-4.012615e+00	9.133078e-05
HB	-1.363059e-01	0.074419	-1.831591e+00	6.883783e-02

An opposite effect was presented in the second equation, in Table 5.3, which was characterized by the negative impact on endogenous variable. It would be interesting to check these relations and the impact of these effects in the following sections. The assessment of the inner model was almost interpretable from results in Table 5.4. "R2" for the endogenous variables, highlights a very slow value for "Financial Performance" whereas a comfortable 0.63 for HB. A slow "R2" value meant that exogenous latent variables had weak power to explain variance of the endogenous latent variable.
The "Block Communality", which measured how much a block variability was reproducible by the latent variable gave us good values. Good outcomes came from "Mean Redundancy" on HB, which meant that "Clients' Risk Aversion" and "Geographical Proximity" could predict more than 50% of the variability of HB whereas a not sufficient result for "Financial Performances" was obtained as it performed a value around 10%. An empirical explanation of this result will be given analysing heterogeneity of the dataset in the next chapter.

TABLE 5.4: Inner Model Assessemnt

	Type	R2	Block Communality	Mean Redundancy	AVE
Clients' Risk Aversion	Ex	0.00000	0.92757	0.000000	0.92757
Geographical Proximity	Ex	0.00000	0.76039	0.000000	0.76039
HB	End	0.63776	0.81271	0.518323	0.81271
Financial Performances	End	0.12093	0.82517	0.099789	0.82517

Effects between latent variables were the most interesting topic in PLS PM analysis; a sum up presentation of direct and indirect effects was presented in Table 5.5.

TABLE 5.5: Effects between latent variables in the Inner Model

Relationships	Direct	Indirect	Total
Clients' Risk Aversion -> Geo Proximity	0.00000000	0.000000000	0.00000000
Clients' Risk Aversion -> HB	0.79189813	0.000000000	0.79189813
Clients' Risk Aversion -> Fin Performances	0.00000000	-0.107940365	-0.10794037
Geo Proximity -> HB	0.03721486	0.000000000	0.03721486
Geo Proximity -> Fin Performances	-0.29861630	-0.005072604	-0.30368891
HB -> Fin Performances	-0.13630587	0.000000000	-0.13630587

A positive direct effects, as theorised in past chapters, were the "Clients' Risk Aversion" and "Geographical Proximity" on "HB". Whereas we obtained a negative direct effects represented by "Geographical Proximity" and "HB" on "Financial Performances", with the first relationship which deviated quite surprisingly from our theory. There were also the presence of two indirect effects, namely "Clients' Risk Aversion" on "Financial Performances" and "Geographical Proximity" on "Financial Performances".

Before drawing conclusions, we should wait the validation which could give an authoritative assessment of the estimated coefficients.

As a first step to give an initial assessment of the model, before starting validation process, was represented by the global index of goodness-of-fit which in our case was equal to 0.5587, highly penalized from the R2 performance of "Financial Performances" estimation. Following the rule of thumb, which not determine model's statistical significance, a very good model should had a GoF around 70% but in our case, we were quietly below this threshold valuing the model as having an adequate prediction power.

5.2.4 Validation

Due to absence of stringent distribution assumptions in PLS-PM, a resampling method was the most indicated one to check the validity of parameters estimated during outer and inner assessments and to measure their variability.

We use the bootstrapping method, which was a non-parametric approach that recursively repeat for 100 times the evaluation of the most important results of the

model. From Table A.3 to Table A.5, on Appendix A, we could find part of the output of the resampling method.
As we could notice, weights and loading did not have zeros in the 5% confidence interval distribution, which meant that the distribution had only positive values. In case of presence of zero in the distribution, we could say that the coefficient of distribution was not significant at 5% confidence interval.
As conclusion of this analysis, we needed to sum up the output and compare the results with our initial thoughts.

5.2.5 Results' Assessment

"Clients' Risk Aversion" had a coherent effect both on "HB" and "Financial Performances", confirming that it could be considered an important cause of home bias and was able to negatively affect branch performances. The link between "Geographical Proximity" and "HB" without resampling highlighted a positive effect which might validate our theoretical idea but, with the resampling method, we saw that zero was in the distribution. We could not confirm this effect with a 5% confidence interval, but it might be confirmed with a less stringent interval such as 10% confidence interval. From our available data, the geographical effect seemed to have a negative impact on branches' performances, as verified from resampling method, deviating from both literature and our theory. This deviation could be interpreted as territorial and social dynamics not considered by the model.
"HB" seemed to have a negative effect on "Financial Performances", this could be the results of MIFID capital requirements which discourage banks to issue a high quantity of CCBs' bond in their clients' portfolios to avoid incurring in sanctions.
In the next chapter, will be presenting a possible motivation for these results, introducing the concept of heterogeneity of the dataset an providing feasible solution.

Chapter 6

Heterogeneity and REBUS PLS

6.1 Heterogeneity

The analysis conducted in this chapter was independent from categorical differentiation that could arise between banks as they belonged to different CCBs. Indeed, we did not categorize branches as an operative part of a single bank but as a single entity in order to detect possible common factors between other branches. In this way, we tried to discover some hidden informations between data, in order to group branches as independent players.

Even though we may though about a model which fitted in a good way data and that its estimated parameters were highly representative of all the individuals of the dataset, it was possible that these assumptions did not hold perfectly and that there was presence of diversity in our dataset which could be express by the concept of heterogeneity. The reason why we made this advanced analysis was that there might be bias in the model studied in Chapter 5 and some conclusions about relationships in the inner model might be affected leading to a possible misinterpretation of the model. The creation of more homogenous classes should improve prediction performance and the interpretability of the model.

During the analysis made on Chapter 5, we dealt with a priori clustering approach assuming homogeneity due to bank classification, but with this new specification we avoiding this categorization and dealt with unobserved heterogeneity.

The easiest way to deal with unobserved heterogeneity was "a two step-procedure, one in which we combined cluster analysis in the first step with a multi-group analysis in the second step"[27]. In essence, we formed groups of data throw clustering analysis and performed multi-group analysis in the separate models for each cluster. An interesting approach to detect unobserved classes inside dataset was REBUS-PLS approach made by Vinzi and Trinchera. Cluster analysis aimed to create classes composed by objects sharing similar behaviors and in our case we used a hierarchical method following the "agglomerative approach". It tended to start the analysis from the bottom, where single branches were the observations and step by step a more agglomerative separated clusters were created, with the extreme case where a unique cluster was obtained. Multi-group analysis aimed to find substantial differences in previously clustered groups, more precisely if there existed differences between their parameter estimates as loadings, weights and path coefficients.

6.1.1 The Rebus-PLS Algorithm

REBUS-PLS, where REBUS stand for Response Based Unit Segmentation in PLS-PM, aimed, thanks to the homonymous algorithm, to detect latent classes in the model computed previously with PLS-PM approach. It could be considered a response-based clustering technique which was inspired by clustering analysis.
The REBUS-PLS algorithm presented by Trinchera in 2007, was recursive algorithm which founded its roots in the original PLS algorithm presented in Chapter 3. The main differences were in the introduction of the closeness measure (CM) between units and model based on residuals. The intrinsic idea was that "if latent classes exist, units belonging to the same latent classes will have similar local models". Furthermore, "if a unit is assigned to the correct latent class, its performance in the local model[1]... will be better than the performance of the same unit considered as supplementary in the other local models"[26]. Another interesting feature of this model was the simultaneous computation for measurement and inner model when clusters were applied. A final measure of goodness for each local model was the GoF, the same measure of fitness used in PLS PM.

[1]The model referring only on a small subset of observation called "classes"

On its first and second non iterative steps, the algorithm computes the estimations for global model, the model estimated on the whole sample, just performed a PLS PM analysis with the computation of communalities and residuals from the inner model. The third phase was characterized by a cluster analysis on the residuals of the global model and the specification on the K classes we wanted to populate with dataset units. Then, the iterative part started with the estimation of K local models "performing a hierarchical cluster analysis on the computed residuals" for both measurement and inner models. When objects were defined, a PLS PM analysis was run for each K class and residuals form K class parameters output and global model parameters leaded to the computation of CM for each unit of the dataset, following a specific complex formula in the algorithm. Furthermore, units were allocated to the model by which performed the smallest CM, for all units until K models where were estimated k classes. All this process was iteratively repeated until, following the rule of thumb, the threshold of approximately 5% of units changing class from an iteration to the next was reached. The overall quality of the REBUS-pls model was assessed by Group Quality Index (GQI) which was inspired by the usual goodness of fit but in this case it referred to multi-group analysis. Indeed, when local models performed a better GQI than the global model's GoF, we could affirm that segmentation created more homogenous classes than the homogeneity presented in the global model.

6.2 Home Bias segmentation and REBUS PLS application

Considered all the analysis made in Chapter 4, with the estimation of the PLS PM model on the whole dataset, we would motivate and got better some of those coefficients which gave us a different interpretation from our initial idea or had a low statistical significance. There might be a sign of the presence of unobserved heterogeneity in the data which could motivate these results.

A first sign of instability of the model were represented by a very low value of R2

and redundancy for "Financial Performance" with the contribution of a quite low GoF (0.5587) should suggest the presence of heterogeneity among dataset and we should look for more homogenous segments.
Before the application of the REBUS PLS algorithm, we chose to split our dataset following the criteria of high or low values of home bias. As discriminant, we decided to us PLS PM standardized scores in Chapter 4, taking values assigned to "HB" latent variable. We defined a branch with a high presence of bond home bias those that presented higher value than the mean of the standardized scores. In our case, the mean value was equal to 0.5907, so all the branches were presenting values lower than this threshold were considered affected by the low presence of home bias. In this way we created a dichotomous variable in our dataset, useful to crate two subsets: "High HB" and "Low HB".
This passage gave us the possibility to see the intrinsic characteristics of both branch profiles and if existed factors which affected more or less in our model.
In the next sections we decided to focus less on classical PLS PM application on the two new datasets, which still followed the previous PLS analysis, in favor to the outcome of the REBUS PLS. The same manifest variable used in Chapter 4 were adopted with new datasets.

6.3 Branches with High Home Bias

We presented the main results of PLS PM, which comprehends the first and second phases of the algorithm, on High Home Bias dataset, characterized by branches which presented high values of this phenomenon. It was easy to see, in Figure 6.1, the inner model together with path coefficients which highlighted the behavior of the model in presence of high home bias. Even though, from a graphical point of view relationships among latent variables seemed changed, all latent variables were significative at 5% confidence interval.
Furthermore, we proceeded with the third phase of the algorithm where we selected

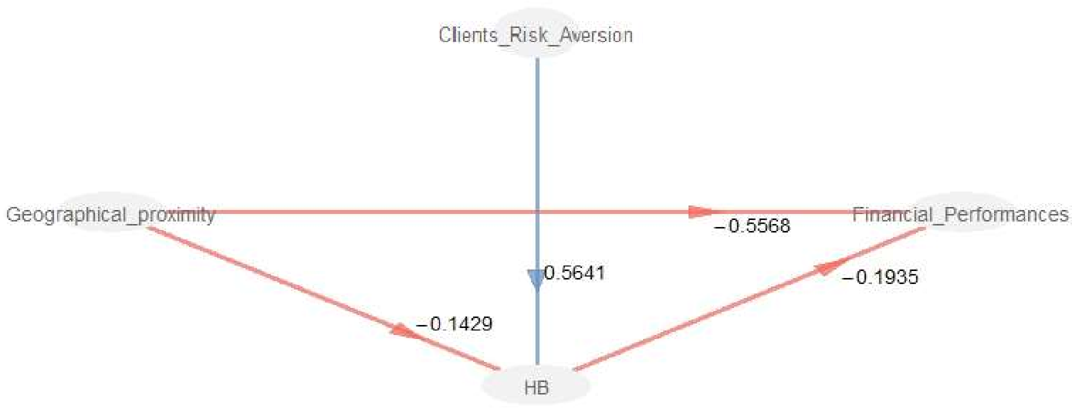

FIGURE 6.1: PLS PM - Inner Model of "High Home Bias" subset

the number of K classes thanks to a cluster dendrogram[2], graphically exposed in Figure 6.2. The number of possible classes we took into consideration may vary from 3 to 4, but since the 3 segmentations model provided a high value of the Group Quality Index (GQI) we choose this last specification.

A first analysis on results Class 1 and Class 3 in Figure B.3, which represented respectively the 17% and 36% of the whole dataset, provided almost all loadings coefficients greater than the ones in the Global model. They also had a very good results on GoF both over 0.70 ,which were higher than the Global model, where was confirmed the presence of higher homogeneity in the classes rather than the heterogeneity in the whole dataset.

Similar conclusions could be expressed regarding Class 2 and Class 4 in Figure B.4, where, excluding some loadings values, they were in line with previous results obtaining a similar GoFs. Generally, we said that model fitted very well the new data segmentations, underling the advantages given by applying the model on homogeneous classes rather than on heterogeneous dataset.

Regarding inner model, main results were presented in Table A.6. Classes 1, 2 and

[2]Typically used in cluster analysis to put in evidence either individual data or clusters in a hierarchical representation

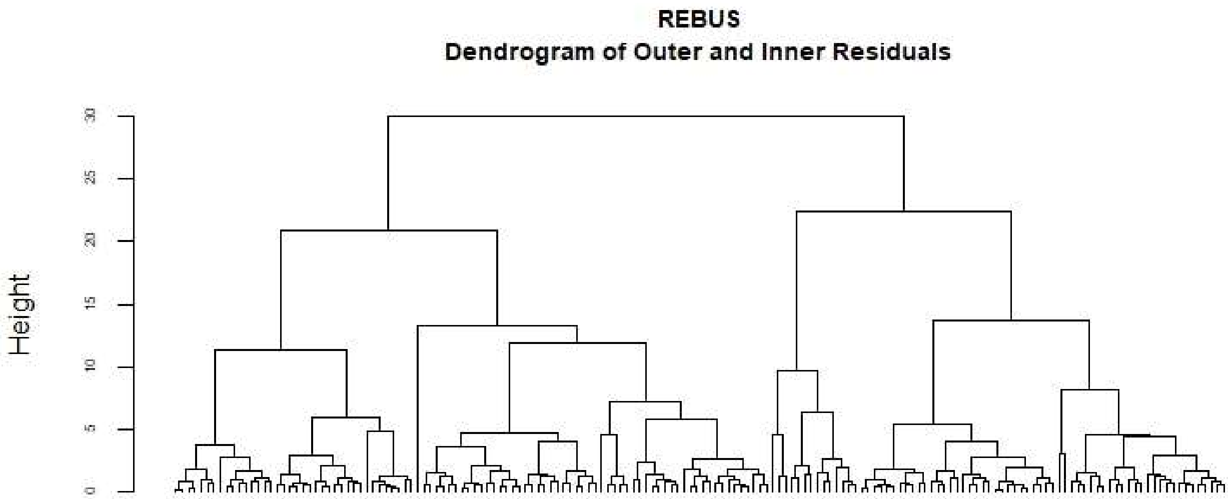

Figure 6.2: Dendrogram of "High Home Bias" subset

3 had almost the same behaviors characterized by negative relationships with all variables, except for "Clients' Risk Aversion" which for all classes had a positive influence on "HB" and in Class 3 geographical Proximity relation and HB had positive sign.

They had almost the same coefficients on average redundancy, whereas for R2 Class 2 seemed to perform better. Class 3 was the most populated one, but presented path coefficients not in line with High Home Bias behavior even though it satisfied all sufficient requirements.

If we should chose the most representative classes of the branches with high home bias, we would chose Class 2 because was the second most populated class, it shared the sign of path coefficients with the majority of the classes, it provided higher loading values than global model and had an important GoF.

6.4 Branches with Low Home Bias

As we could notice in Figure 6.3, the inner model was shown together with path coefficients which highlighted the behavior of the model in presence of low home

TABLE 6.1: Path coefficients and quality analysis: High Home Bias

	Global	Class.1	Class.2	Class.3	Class.4
Number of Units	95	16	27	34	18
Proportions	100%	17%	28%	36%	19%
Path Coefficients					
Clients' Risk Aversion->HB	0.6218	0.8933	0.7348	0.4227	0.6573
Geo Proxi->HB	-0.0982	-0.1487	-0.2349	-0.5207	0.2015
Geo Proxi->Fin Perf	-0.5307	-0.6629	-0.8701	0.8117	-0.7957
HB->Fin Perf	-0.2368	-0.2233	-0.0745	0.0203	-0.1108
Aver.Redu					
Red.HB	0.1560	0.6656	0.3984	0.3643	0.2980
Red.Fin Perf	0.8487	0.3163	0.6726	0.6075	0.6245
R2					
R2.HB	0.4019	0.9234	0.6448	0.4806	0.5486
R2.Fin Perf	0.3297	0.3425	0.7184	0.6411	0.7142
GoF	0.5164	0.7377	0.7501	0.7014	0.7169

bias. Considering we were dealing with low Home Bias branches, some differences between the previous model were expected. Indeed, the positive graphical relationships between latent variables was validated by bootstrap process, except for the effect of "Geographical Proximity" to "Financial Performances" which presented a zero in the distribution forcing us to affirm the non significativity at 5%.
Moved on local models estimation we had as reference Figure 6.4 which helped us to select 3 classes to start the REBUS PLS algorithm. The choice was also supported by a high value of GQI, around 0.70, which was the highest among numbers around 3 and attested the quality of numbers chosen.
The total number of branches in this subset was 71, and over a half of the individuals were represented by Class 2 with 55% of the global model. All the classes, as we could see from Table A.7, provided higher loadings with respect to Global Model but also the results for communalities and mean redundancy were definitely better. R2 for both endogenous variables were higher and suggested that clustering techniques applied to the dataset, with the intention to deal with homogenous classes, were correct. Advantages were also highlighted form high Goodness of Fit indicators which were greater than the one provided by Global Model. Also in this subset, we

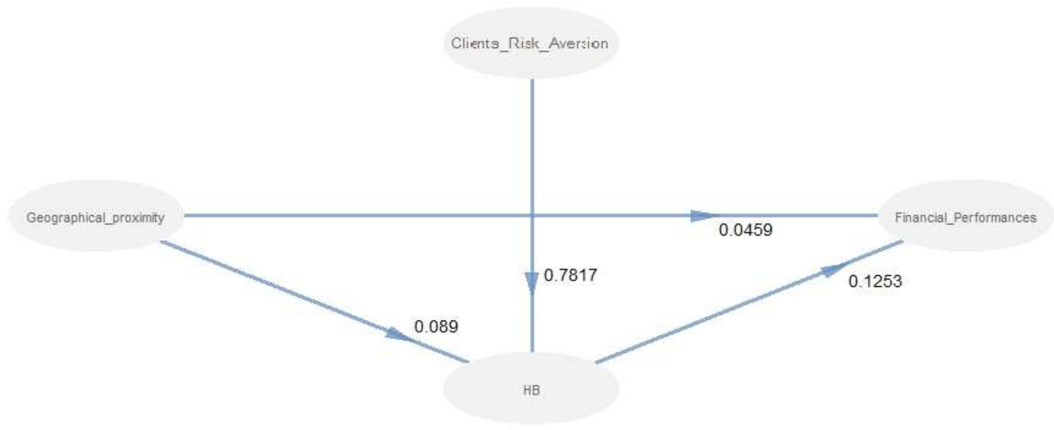

FIGURE 6.3: PLS PM - Inner Model of "Low Home Bias" subset

should choose Class 2 as most representative of the dataset due to numerosity of the class, goodness of loadings and weights estimated. In addition, selecting a Class with a very high GoF, like Class 3 (0.8827), might influence our thoughts about the model and leading to misrepresentation of future analysis.

6.5 Analysis of Branches belonging to High or Low Home Bias

After the previous analysis, we selected one model for each subset as considered the "best models". With the term "best models", we meant that branches inside the subset constituted a group of homogeneous branches which whom the core element of distinction between the other subset was the level of home bias.
Essentially, we could say that both Classes 2 for High Home Bias subset and Low Home Bias subset were the best proxy of branches sharing similar policies obtaining similar results. They could be considered as "best models" due to the high presence

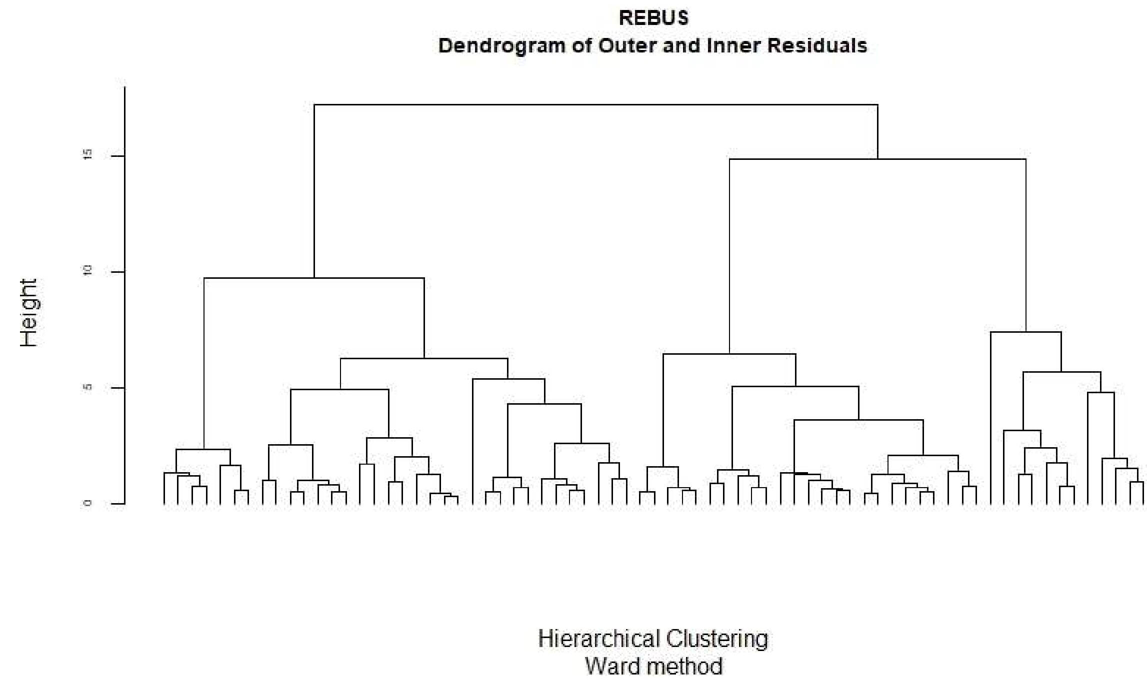

FIGURE 6.4: Dendrogram of "Low Home Bias" subset

of unit in these classes, affirming that these branches represented together valid proxies for respective subsets

Thanks to the REBUS segmentation, we were able to detect which of the branches studies have similar characteristics with respect to the variable analysed. It should be interesting to see how many similarities branches have, even though belonging to a different mother-bank.

Clusters were created to sum up branches having similarities in all variables studies in the previous chapter, so we can say that if two branches can be considered similar, they should share the same modus operandi and same policies in running a bank. Looking at clusters and considering the assumption of same policies, we make come conclusions together with cluster effects.

The selection of the "best model" in the subset was just made to allow us in the discovery of the branches inside each single branch, and verify if their similarities were drive from specific variables and how they differs in relation to the level of home bias.

TABLE 6.2: Path coefficients and quality analysis: Low Home Bias

	Global	Class.1	Class.2	Class.3
Number of Units	71	15	39	17
Proportions	100%	21%	55%	24%
Path Coeffcients				
Clients' Risk Aversion->HB	0.7817	0.4686	0.8093	0.8100
Geo Prox->HB	0.0890	-0.4548	0.1849	0.2600
Geo Prox->Fin Perf	0.0570	0.8482	0.5317	-0.9526
HB->Fin Perf	0.12533	0.8750	-0.5031	0.1025
Aver.Redu				
Red.HB	0.4946	0.4101	0.4836	0.8297
Red.Fin Perf	0.0203	0.5317	0.3726	0.7683
R2				
R2.HB	0.6525	0.5817	0.7224	0.9133
R2.Fin Perf	0.0210	0.5565	0.3887	0.7959
GoF	0.5204	0.6755	0.6630	0.8826

6.5.1 Elements of classes

We wanted to investigate which branches were in each subset and highlighted the main characteristics.

In lights of the results presented in Table 6.3 as the summary elements of branches affected by high or low presence of home bias some consideration could be made.

"Clients' Risk Aversion" seemed to affect in both cases in a positive manner home bias, confirming our theoretical idea assessing it was one of the most important source of this phenomenon.

"Geographical Proximity" had a negative incidence in HB in the High HB subset whereas a positive one in the other subset. Seemed that an increment of the closeness of the bank to their clients, geographically and physically speaking, led Home Bias to decrease and the same negative effect was referred to Financial Performances. A possible interpretation was represented by the presence of high levels of home bias in banks belonging to High HB subset. There might existed a possible adequate level where, once exceeded, the marginal increment of geographical proximity would resulted to generate negative effect. For "Financial Performances"in High HB, the increase of "Geographical Proximity" was considered as a possible

TABLE 6.3: Path coefficients and quality analysis: Comparison High and Low Home Bias

	High HB Proxy	Low HB Proxy
Number of Units	27	39
Path Coefficients		
Clients' Risk Aversion->HB	0.7348	0.893
Geo Prox->HB	-0.2349	0.1849
Geo Prox->Fin Perf	-0.8701	0.5317
HB->Fin Perf	-0.0745	-0.5031
AverRedu		
Red.HB	0.3984	0.4836
Red.Fin Perf	0.6726	0.3726
R2		
R2.HB	0.6448	0.7224
R2.Fin Perf	0.7184	0.3887
GoF	0.7501	0.6630

costs sustained from bank to expand its own presence on the territory and would not have a marginal effect caused from the already high level of proximity. In light of these results, we thought at "Geographical Proximity" as a nonlinear function with respect to "HB" and "Financial Performances" where, reached a certain adequate level of "proximity", benefits on "HB" and "Financial Performances" were zero or negatives. With respect to the subset "Low Home Bias", seemed that the level previously cited was not reached and banks might increment both presence on the territory and personal to increase HB and obtain better financial performances. A more detailed analysis, referred to bank classification was conducted.
In Figure 6.5, was exposed how High HB and Low HB subsets how banks categories were distributed. Knowing that High HB subset was composed in majority of branches belonging to BCDMR and CENMR we could say that in these banks, generally, the incidence of "Home Bias" phenomenon on "Financial Performances" has a negative effect quantified around -0.0745. So, when home bias increased of 1%, Financial performances were used to decay of over a 8%.
Regarding the other subset, the relationship still remained negative, but the incidence was much more marked as it was equal to -0.5031. These results should partially confirm our starting idea, where presence of high home bias would lead to a better financial performance. In this case, we said that it was a negative result but

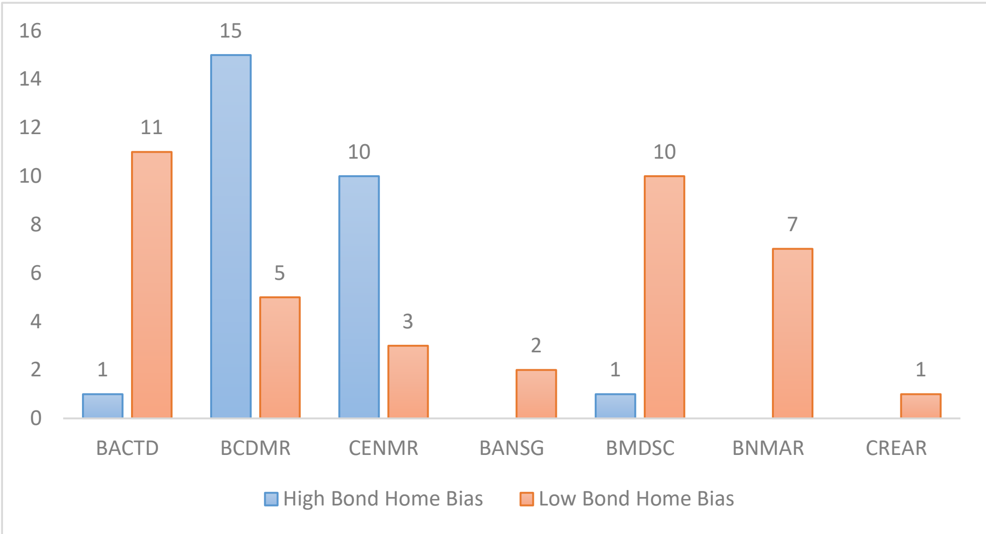

FIGURE 6.5: Graphical comparison between High Bond Home Bias branches and High Bond Home Bias branches

much better that the damage occurred to the banks without the presence of home bias. Results were in line with the analysis made on banks' asset allocation and performances, as BCDMR and CENMR issued in 2014 high quantitative of own bonds in clients' portfolio reaching respectively 45.74% and 46,47% incidence the total portfolio. We supposed that these banks on 2014 actuated a policy of self-financing, issued bonds to their clients and obtained a better result with respect to the theoretical one they would achieve in the case of low issuing of bonds. An important implication should be made with respect to direct fund raising. Banks which belonged to High HB subset presented a marked increase of direct funding, showed in Figure 4.4, like the path made by BCDMR form 2012 to 2014, whereas banks belonged to Low HB, like BNMAR had a decreasing trend line. On these results, we could made the hypothesis that banks in High HB subset might actuated a policy directed to the increase of the self-financing inserting almost 50% of CCB's issued bond in clients portfolios. This should be eventually assessed from authorities as non-ethic conduction of direct raising of funds and negligence of financial advisors

for those banks which presented similar behaviors.
Empirical results did not completely confirm our thesis, but we justified this assuming that the negative financial results of banks affected by home bias could be the effects of environmental and socio-political dynamics we did not catch in our model due to lack of information and data. For banks belonging to High HB subset, the high level of home bias, caused from a high presence of CCBs-issued bonds in clients portfolios, resulted to be financially advantageous and a cheap way to self-finance their activities.

Chapter 7

Mediator and Moderator Effects

7.1 Difference between mediation and moderation

In this chapter, we went through the analysis of possible different interaction studied so far. Two examples of them were represented by the concepts of mediation and moderation. The former had the "function of a third variable, which represents the generative mechanism through which the focal independent variable was able to influence the dependent variable" whereas the latter was represented by a variable, which was split "into subgroups that establish its domains of maximal effectiveness in regard to a given dependent variable"[28].

7.2 Moderated Regression Analysis

In Chapter 5, we considered manifests variables as group of branches where observable characteristics such as number of clients, average client portfolio composition and many others that were grouped in a so called "segmentation variables", synthesized relationships between latent variables. Due to this particular technique, a problem concerning heterogeneity could arise when there was the suspicion that estimates of the model might change the meaning of the study. This was attributable to the presence of subgroups between manifest variables, which taken singularly might change the total effect of a variable on another. In this way, it should be easier to understand the total effect and if there might exist a specific group of manifest variable that affect the relationship. This technique led us to the creation of "moderator variables" which could modify both the strength and direction of a relationship between latent variables.

Researchers were used to use moderator variables to make multi-group analysis, taking advantage of "transforming the continuous variable into a categorical variable" following the approach of dichomotization.

We could imagine to create a moderator and insert them in the model as latent variables.

The moderators could affect "the form of strength of a relationship between an independent variable and a dependent variable"[28] where the moderator "interacted to reflect that the effect of independent variable on the dependent variable depended on the level of the moderator"[29].

7.3 The Categorical Variable Approach

We created two different analysis, one on the effects between Geographical Proximity and Home Bias and the other between Clients' Risk Aversion and Home Bias. We run this type of analysis due to the nature of the variables as they can be considered categorical variables. It was possible thanks to the a priori clustering represented by the natural belonging of branches to different mother-banks.

Our main goal was to find the contribution on the relationship of different bank, understanding if the cause-effect link was driven by some banks instead of others.

In order to make a proper study of the single contribution effect we created dummy indicators. To do that, we took the number of bank -1 taking the 6th bank "CENMR" as reference group[1], creating dummy variables for the other 7 banks, assigning a value "1" when the informative variable was referring to a specific bank and "0" in all other cases. Dummy variables had all the same length equal to the number of original statistical units.

[1] We took the number of banks -1 to avoid multicollinearity problems. Reference group was selected in an arbitrary way

No clustering was applied during this study, the dataset of reference was the same used in Chapter 5.

7.3.1 Geographical Proximity and Home Bias Relationship

The PLS PM model was run with 7 dummy variables representing all bank's geographical mediating effects and 7 interaction variables as the product between dummies and "Geographical Proximity" as independent variable and "HB" as the dependent variable.
We directly reported on Table 7.1 the results with the corresponding output coming form the bootstrap validation process.

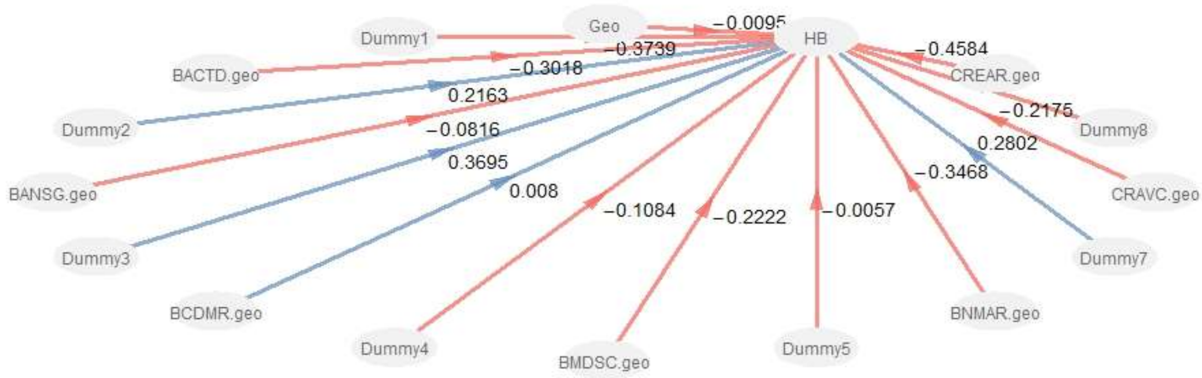

Figure 7.1: Moderators in the relationship between Geographical Proximity and Home Bias

We analysed the moderating effect through significance validation for interaction variables. In this way, we were able to assess which bank leaded the relationship between "Geographical Proximity" and "HB". Those banks, which showed significant effect might be considered banks where the relevance of be "close to clients" predicted better than other banks the home bias phenomenon.

It was important to have a graphical representation of the model on Figure 7.1 to have an overall idea of the dynamics. We could have made some conclusions only thanks to the bootstrap validation process which had the duty to clarify which of the selected effect in the model had statistical relevance as moderating effect. It was important to say that the moderator hypothesis was supported if the interaction between dummy and Geographical Proximity was statistically significant and that coefficient values were computed with respect to the omitted bank "CENMR", used as reference. As we saw from Table 7.1, we had only three bank's variable coefficients which did not include zero in the confidence intervals, namely: "BACTD.geo", "CRAVC.geo" and "CREAR.geo" whereas all the other interaction variables could be considered not statistically significant. In mean, the three previously cited banks moderated/interacted negatively the existing relationship between Geographical Proximity and HB.
This kind of heterogeneity was due to the fact that relationship between latent variables was not constant, but rather depended on moderators' value which were different banks with different policies and behaviors.

TABLE 7.1: Moderators effects - Geographical Proximity to Home Bias

	Original	Mean.Boot	Std.Error	perc.025	perc.975
Geo -> HB	-0.0795	-0.0223	0.0833	-0.2113	-0.0232
Dummy1 -> HB	-0.3739	-0.4109	0.1940	-0.6494	-0.1683
BACTD.geo -> HB	-0.2222	-0.2161	0.1266	-0.3885	-0.0530
Dummy2 -> HB	-0.3018	-0.4475	0.2289	-0.8018	-0.1227
BANSG.geo -> HB	-0.0057	0.0482	0.2041	-0.2617	0.3624
Dummy3 -> HB	0.2163	0.1579	0.1515	-0.0368	0.4383
BCDMR.geo -> HB	-0.3468	-0.0099	0.3999	-0.4478	0.4992
Dummy4 -> HB	-0.0816	-0.1968	0.2055	-0.4529	0.1451
BMDSC.geo -> HB	0.2802	0.1393	0.3100	-0.4850	0.4152
Dummy5 -> HB	0.3695	0.2729	0.1691	0.0379	0.5343
BNMAR.geo -> HB	-0.2175	-0.0923	0.3488	-0.6221	0.3792
Dummy7 -> HB	0.0080	-0.0389	0.0790	-0.1117	0.1205
CRAVC.geo -> HB	-0.3786	-0.3122	0.1783	-0.5437	-0.0676
Dummy8 -> HB	-0.1084	-0.1360	0.2139	-0.4308	0.1896
CREAR.geo -> HB	-0.4584	-0.5522	0.2574	-1.0815	-0.2945

Negative coefficient from moderator of BACTD to HB were interpreted as follows:

took as reference level of geographical proximity the value of BACTD, the relationship between Geographical Proximity and HB was equal to -0.0795. If the Geographical Proximity became higher, an increase of one standard deviation point, implied that relationship between Geographical Proximity and Home Bias decreased by the size of the interaction term, obtaining the value of -0.0795 - 0.2222 = -0.3017. When BACTD geographical proximity increased by more than a standard deviation point, Geographical Proximity became even more negatively significant for the explanation of HB. The same reasoning could be made for all the other banks, which presented statistically significant effects.

7.3.2 Client's Risk Aversion and Home Bias Relationship

A second moderator effect might arose in the relationship between Clients' Risk Aversion and Home Bias. As a first step, we run the PLS obtaining a graphical representation showed in Figure B.5 and the results together with the bootstrap validation test in Table A.8.
With this latent variable, we saw the absence of moderating effect which were considered statistically significant at 5%. Indeed, there were zeros in coefficient bootstrap distribution which allowed us to affirm that no significant increase or decrease in interaction variables' level affected the relationship of CRA on HB.

7.4 Mediator effect

As we saw from previous chapters, the unique indirect relation of the model was represented by Client's Risk Aversion to Financial Performance. The idea was to test this relationship among two clusters we have made before, and check is HB played the role of mediator in the models.
In order to test this mediating effect, we made a three step analysis: first we checked the relationships between Client's Risk Aversion and Financial Performances, then we estimated the relationships between Client's Risk Aversion and HB and as a last step we verified the link between HB and Financial Performances.

7.4.1 High Home bias

A summary of the three effects studied above was presented in Table 7.3. The direct effect of Clients' Risk Aversion on Financial Performances was negative, but not significant at 5%. In this way, we thought that the effect of Clients' Risk Aversion was firstly transferred to HB and then, using HB as mediator, the effect impacted on Financial Performances.

TABLE 7.2: Path coefficients for mediation effect on High Home bias subset

	Original	M.Boot	Std.Error	perc.025	perc.975
Clients' Risk Av.-> HB	0.5803	0.5832	0.1826	0.1872	0.7989
HB -> Financial Perf	-0.1655	-0.2383	0.06670	-0.3662	-0.1570
Clients' Risk Av.-> Financial Perf	-0.1548	-0.1610	0.1039	-0.2781	0.1401

These requirements were satisfied as other two effects under analysis were significant at 5%, as could be seen from absence of zero in the bootstrap distribution. A graphical representation was also presented in Figure 7.2.

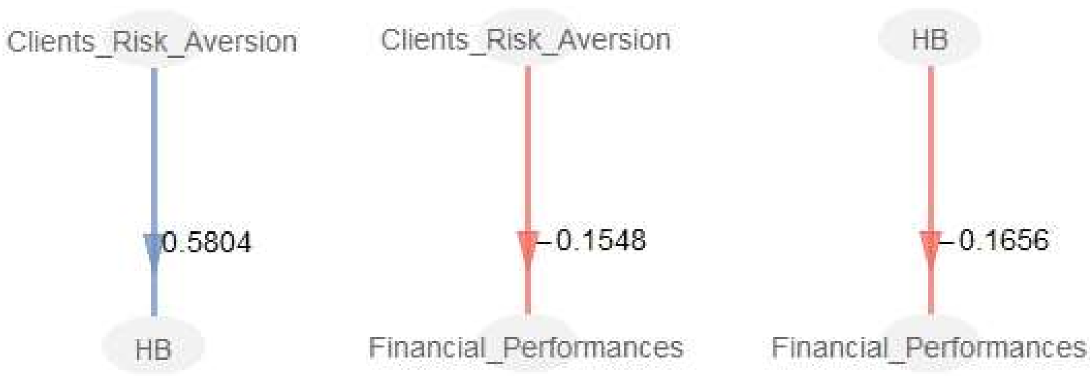

FIGURE 7.2: Path Coefficient representation for mediation effect - High Home bias subset

We could say that HB played the role of a mediator but to validate these results we used the Sobel Test. Form Table 7.3, we saw a p-value which allowed us to reject the null hypothesis of indirect effect equal to 0. There existed a significant indirect effect of Clients' Risk Aversion on Financial Performances related to the subset of banks affected by high home bias.
Similar tests were run to validate this result, namely Aroian Test and Goodman Test[2].

TABLE 7.3: Validation test on indirect effect - High Home Bias subset

Test	Test Statistic	Std. Error	P-value
Sobel Test	-1.98747	0.04832	0.04687
Aroian Test	-1.93203	0.04970	0.05335
Goodman Test	-2.04797	0.04689	0.04056

7.4.2 Low Home bias

It was interesting to check if Clients' Risk Aversion had the same indirect effect through mediation on Financial Performances with respect of the subset which presented low levels of home bias. The procedure followed the one used for the other subset and the presentation of path models between latent variables in Low Home Bias case was presented in Table 7.5.

TABLE 7.4: Path coefficients for mediation effect on Low Home bias subset

	Original	M.Boot	Std.Error	perc.025	perc.975
Clients' Risk Av.-> HB	0.8492	0.8528	0.01891	0.8226	0.8771
HB -> Financial Perf	-0.3728	-0.1056	0.3087	-0.4536	0.0125
Clients' Risk Av.-> Financial Perf	-0.2468	-0.065	0.2218	-0.3214	0.3245

The model under consideration highlighted a similar behavior to the one in the High Home Bias subset; as we saw, there existed a significant positive link between

[2]The Sobel Model is an approximate significative test for the indirect effect of the independent variable on the dependent variable via mediator. The Sobel test used coefficient between independent variable to mediator variable and the one form mediator to the dependent variable with standard error of the coefficients. The Aroian add also a third term of variance in the formula's denominator, whereas theGoodman subtract it.

Client Risk aversion and HB and a slightly[3] significant negative one between HB and Financial Performances.

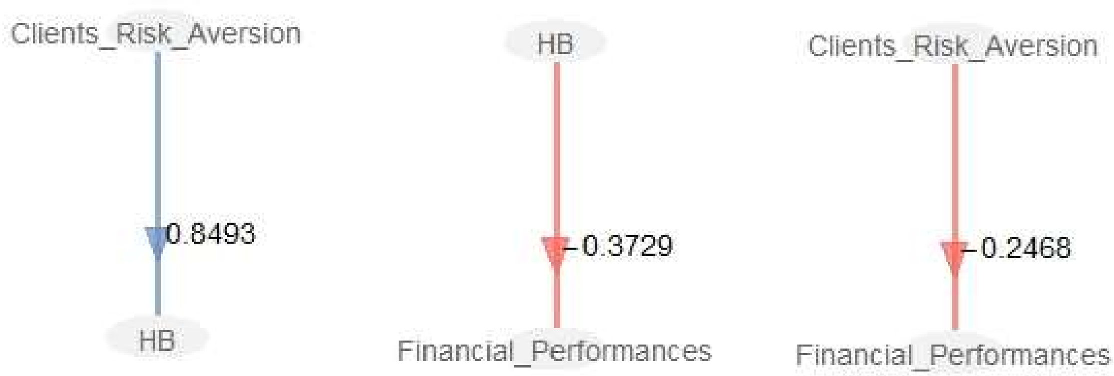

FIGURE 7.3: Path Coefficient representation for mediation effect - Low Home bias subset

Figure 7.3 represent singularly relations between latent variables.
An assessment of path coefficient expressed was given by results presented in Table 7.5 coming from the validation tests on indirect effects.

TABLE 7.5: Validation test on indirect effect - High Home Bias subset

Test	Test Statistic	Std. Error	P-value
Sobel Test	-1.2072	0.2622	0.22735
Aroian Test	-1.2069	0.2623	0.22747
Goodman Test	-1.2075	0.2621	0.22724

As a matter of fact, hypothesis of presence of indirect effect from Clients' Risk Aversion to Financial Performances was rejected from all three tests.

[3]The presence of a zero in bootstrap distributions should lead us to reject the significativity of the path coefficient at 5% confidence interval but despite that we could affirm that the right tail of the distribution was populated by units very close to be negative. We were also likely to accept the 5% confidence interval because even though bootstrap represented one of the best validation methods, it remained a recursive estimation where the distribution change its shape every time computation was performed. This was probably the case where the distribution exceeded on positive values creating suspicious of coefficient significance.

As a final result, we obtained that Clients' Risk Aversion effect on Financial Performances was significant only in the subset composed of banks with high levels of home bias. It seemed that the indirect effect was well transferred through HB as mediator which attested that Clients' Risk Aversion affect the negatively Financial Performance. Results partially justified our analysis because there was a presence of indirect effect but we didn't confirm if the impact on dependent variable Financial Performances with a significant result for Low Home bias subset. In that case, we should assess if the indirect effect affected more or less the final results. As we mentioned in Chapter 6, the negative effect of HB on Financial Performances, with or without highlighting indirect effects, was attributable to industry and socio-political dynamics we did not capture in our model, due to lack of data.

Chapter 8

Conclusions

The main topics studied in this thesis were related to analyse the principal source of bond home bias in cooperative credit banks and how they were related with bank's financial performances. Our initial theory was represented from the idea that CCBs were encouraged to sell CCB-issuing bonds to their clients in order to self-financing their own activities, preferring high levels of bond home bias. From our dataset, composed by 8 CCBs based in northeast of Italy, we found that Client's Risk Aversion and Geographical Proximity were positively related to Home Bias, confirming our theoretical thoughts. A deviation form our expectations was represented by the negative relationship between Geographical Proximity and Financial Performances, which affirmed that more presence of the bank, in terms of number of employees, number of financial advisor and number of offices, did not turn into a better economic performance. The other important deviation was represented by the negative link between Home Bias and Financial Performances.
The deviation of empirical results from theory created a suspicious to be dealing with a dataset affected by heterogeneity. First of all, relying of previous results, we divided the datasets in two sub-groups where discriminant was the high or low level of Home Bias. On these new sub-sets we applied PLS PM Rebus algorithm which created groups of homogeneous units. We selected the most representative group of units for each sub-set and we studied the relationship between latent variables. The main finding was that those branches presenting high levels of home bias, even though the relationship with Financial Performances was still negative, had a better results than those with low Home Bias, confirming our starting theory. We found a

strictly relationship between banks having high home bias, obtaining financial advantages and characterized by high percentage of direct fund raising which might thought about a negligent conduction of the branch, maybe damaging clients.
On our last analysis, we found a moderating effect in the relationship between Geographical proximity to Home Bias which amplify the strength of the negative link between variables due to a simultaneous significance presence of geographical effect in some banks, namely BACTD, CRAVC and CREAR. No moderator effect was found in the effect form Clients' Risk Aversion and Home Bias.
Home Bias was attested to represent the mediation effect of Clients' Risk Aversion and Financial Performances only in the subset characterized by brahces with high home bias values. In essence, the mediating effect putted in evidence that this indirect effect could be transferred through Home Bias to Financial Performances only when there existed a relevant presence of CCBs bond in clients' portfolios.
We strongly believed that with a higher number of Financial Performances data related to each branch and a more availability of information to use as manifest variables, the model would provided more exhaustive and trustable results.
It should be interesting to follow this still quite an unexplored field of study, focusing more on branches and CCBs interactions and trying to use PLS PM models which would incorporate more restrictions in order to obtain a higher and trustable informative outputs.

Appendix A

Appendix

A.1 Tables of Chapter 4

Table A.1: Asset Allocation 2014 client's portfolios

	Bcc.Bond	Bond	Gov.bond	Stocks	Funds	Insurance	Derivatives	PCT	CD	Others
BACTD	18,44%	8,04%	32,36%	2,90%	3,37%	0,46%	0,08%	3,51%	29,22%	1,57%
BANSG	44,15%	6,78%	3,83%	6,25%	27,63%	10,02%	0,05%	0,00%	1,24%	0,00%
BCDMR	45,74%	7,85%	7,61%	2,98%	8,56%	24,95%	0,002%	0,00%	2,26%	0,00%
BMDSC	40,08%	3,47%	9,19%	2,12%	23,06%	7,47%	0,01%	0,00%	14,57%	0,00%
BNMAR	52,45%	8,83%	17,53%	2,39%	7,44%	9,29%	0,00%	1,23%	0,79%	0,01%
CENMR	46,48%	7,94%	14,33%	3,55%	14,02%	12,79%	0,005%	0,10%	0,68%	0,05%
CRAVC	43,70%	2,80%	8,39%	3,25%	9,31%	11,66%	0,003%	0,00%	20,69%	0,16%
CREAR	36,06%	11,19%	6,80%	4,55%	30,10%	6,79%	0,11%	0,00%	4,36%	0,00%

Table A.2: Home bias across banks form 2011 to 2015

	2011	2012	2013	2014	2015
BACTD	0,7250	0,6508	0,5515	0,2816	0,0498
BANSG	0,7172	0,7730	0,7817	0,7971	0,8199
BCDMR	0,5418	0,4887	0,6534	0,7227	0,7044
BMDSC	NA	0,6707	0,5989	0,5937	0,6436
BNMAR	NA	0,6494	0,5891	0,5937	0,6030
CENMR	0,7045	0,6820	0,6943	0,6450	0,6924
CRAVC	0,7864	0,7624	0,7766	0,5910	0,5763
CREAR	0,6243	0,5726	0,6641	0,5961	0,6612

A.2 Tables of Chapter 5

A.2.1 Bootstrap Validation Outputs

TABLE A.3: Weights Bootstrap Validation

	Original	Mean.Boot	Std.Error	perc.025	perc.975
Clients' Risk Aversion-per.bbc	0.6030987	0.6016553	0.02691790	0.5621035	0.6769811
Clients' Risk Aversion-per.low.risk	0.4331665	0.4334044	0.01647603	0.3940085	0.4569570
Geo Proximity-Num.tot.risorse	0.2431195	0.1967292	0.18014726	-0.2846000	0.4542256
Geo Proximity-Num.risorse.investimenti	0.4497495	0.5193259	0.17779039	0.3064778	0.9614900
Geo Proximity-branch nclient	0.4575424	0.4106697	0.15075200	0.0554444	0.7556794
HB-HB.Branch	0.4172957	0.4164119	0.01794604	0.3841263	0.4557428
HB-bbc.on.mkt.val	0.4323433	0.4324664	0.02416467	0.3910140	0.4746432
HB-bcc.1	0.2449673	0.2444339	0.02476244	0.1947021	0.2820786
Fin Performances-ROE	0.3083792	0.3230609	0.04592711	0.2411735	0.4005703
Fin Performances-ROA	0.2660261	0.2643584	0.05848714	0.1447913	0.3625914
Fin Performances-Netincome	0.5268068	0.5019482	0.09076942	0.3581102	0.6711504

TABLE A.4: Loadings Bootstrap Validation

	Original	Mean.Boot	Std.Error	perc.025	perc.975
Clients' Risk Aversion-per.bbc	0.9750	0.9763	0.006381	0.9641	0.9863
Clients' Risk Aversion-per.low.risk	0.9510	0.951	0.02149	0.8863	0.9782
Geo Proximity-Num.tot.risorse	0.8867	0.8461	0.1059	0.5696	0.9531
Geo Proximity-Num.risorse.investimenti	0.8533	0.8742	0.05484	0.7895	0.9833
Geo Proximity-branch nclient	0.8755	0.8314	0.10372	0.51224	0.9226
HB-HB.Branch	0.9468	0.9470	0.007643	0.9299	0.960
HB-bbc.on.mkt.val	0.9384	0.9387	0.00807	0.9210	0.9532
HB-bcc.1	0.8129	0.8134	0.036443	0.7382	0.8749
Fin Performances-ROE	0.9029	0.9139	0.026876	0.8673	0.959
Fin Performances-ROA	0.9141	0.9223	0.023975	0.8727	0.9610
Fin Performances-Netincome	0.9080	0.9094	0.02843	0.853	0.9641

TABLE A.5: Effects Bootstrap Validation

	Original	Mean.Boot	Std.Error	perc.025	perc.975
Clients' Risk Aversion -> Geo Proximity	0.0000	0.0000	0.0000	0.0000	0.0000
Clients' Risk Aversion -> HB	0.7918	0.7931	0.0427	0.7050	0.8566
Clients' Risk Aversion-> Finl Performances	-0.1079	-0.1118	0.0386	-0.1793	-0.0389
Geo Proximity -> HB	0.0372	0.0433	0.0374	-0.0340	0.1108
Geo Proximity -> Fin Performances	-0.3036	-0.2828	0.1270	-0.52441848	-0.0521
HB -> Fin Performances	-0.1363	-0.1410	0.0487	-0.2157	-0.0490

A.3 Tables of Chapter 6

A.3.1 Local Models - High Home Bias

TABLE A.6: Path coefficients and quality analysis

	Global	Class.1	Class.2	Class.3	Class.4
Number of Units	166	16	27	34	18
Proportions	100%	17%	28%	36%	19%
per.bbc	0.9871	0.9865	0.9892	0.9955	0.9739
per.low.risk	0.9822	0.9851	0.9880	0.9932	0.9677
Num.tot.risorse	0.9655	0.9767	0.9687	0.9339	0.9792
Num.risorse.investimenti	0.8142	0.8754	0.6974	0.8348	0.9285
branch nclient	0.8807	0.8651	0.9427	0.9346	0.9306
HB.Branch	0.5686	0.7177	0.8361	0.8458	0.5874
bbc.on.mkt.val	0.9626	0.9626	0.7325	0.8948	0.8611
bcc.1	0.9201	0.9623	0.9422	0.9200	0.9328
ROE	0.9398	0.9634	0.9771	0.9772	0.9403
ROA	0.9514	0.9748	0.9734	0.9716	0.9640
Netincome	0.8698	0.9448	0.9522	0.9715	0.9000
GoF	0.5587	0.7376	0.7502	0.7014	0.7169

A.3.2 Local Models - Low Home Bias

TABLE A.7: Loadings comparison between global model and 4 local models

	Global	Class.1	Class.2	Class.3
Number of Units	61	15	39	17
Proportions	100%	21%	55%	24%
per.bbc	0.9608	0.9856	0.9501	0.9696
per.low.risk	0.8455	0.9097	0.8240	0.9564
Num.tot.risorse	0.8165	0.8745	0.8670	0.9594
Num.risorse.investimenti	0.8713	0.6863	0.8223	0.8901
branch nclient	0.8759	0.8401	0.9003	0.9071
HB.Branch	0.9228	0.8800	0.9183	0.9567
bbc.on.mkt.val	0.9466	0.9302	0.9356	0.9755
bcc.1	0.7255	0.6895	0.5386	0.9267
ROE	0.9807	0.9802	0.9769	0.9855
ROA	0.9866	0.9746	0.9813	0.9794
Netincome	0.8563	0.9012	0.8823	0.9234
GoF	0.52037	0.6755	0.6630	0.8826

A.4 Tables of Chapter 7

A.4.1 Inner Model Assessments with Moderating Variables : Clients Risk Aversion - HB

TABLE A.8: Home bias across banks form 2011 to 2015

	Original	Mean.Boot	Std.Error	perc.025	perc.975
Clients Risk Aversion -> HB	0.5967	0.8013	0.4003	0.3693	1.5030
Dummy1 -> HB	-0.4171	-0.1603	0.5373	-0.7392	0.7987
BACTD.cra -> HB	0.1893	0.0259	0.3406	-0.5812	0.3941
Dummy2 -> HB	-0.4093	-0.3540	0.5653	-0.9999	0.6723
BANSG.cra -> HB	0.5603	0.4934	0.5401	-0.5038	1.0639
Dummy3 -> HB	-0.0351	0.2823	0.6604	-0.3546	1.4297
BCDMR.cra -> HB	0.0468	-0.2901	0.6631	-1.4156	0.3648
Dummy4 -> HB	0.0175	0.2675	0.4727	-0.2769	1.0032
BMDSC.cra -> HB	0.0305	-0.1859	0.4040	-0.8025	0.2559
Dummy5 -> HB	-0.0110	0.3668	0.4991	-0.1901	1.2875
BNMAR.cra -> HB	-0.0546	0.4966	0.5174	-0.2305	1.2592
Dummy7 -> HB	0.3683	0.4883	0.4174	-0.0017	1.1670
CRAVC.cra -> HB	0.4280	0.0175	0.8326	-1.4845	0.9857
Dummy8 -> HB	-0.2563	-0.0871	0.9332	-1.5662	1.2650
CREAR.cra -> HB	0.0942	-0.9542	1.2800	-2.4330	1.4214

Appendix B

Appendix

B.1 Figures of Chapter 6

FIGURE B.1: Local Model: Classes 1 and 3

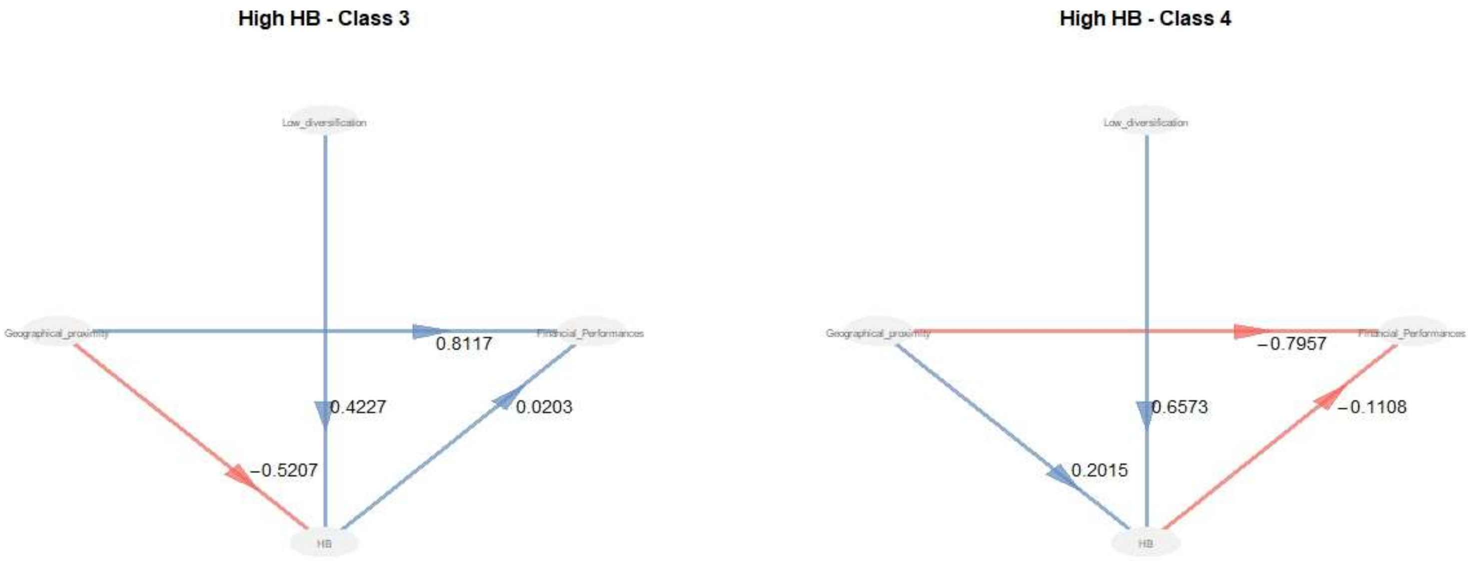

FIGURE B.2: Local Model: Classes 2 and 4

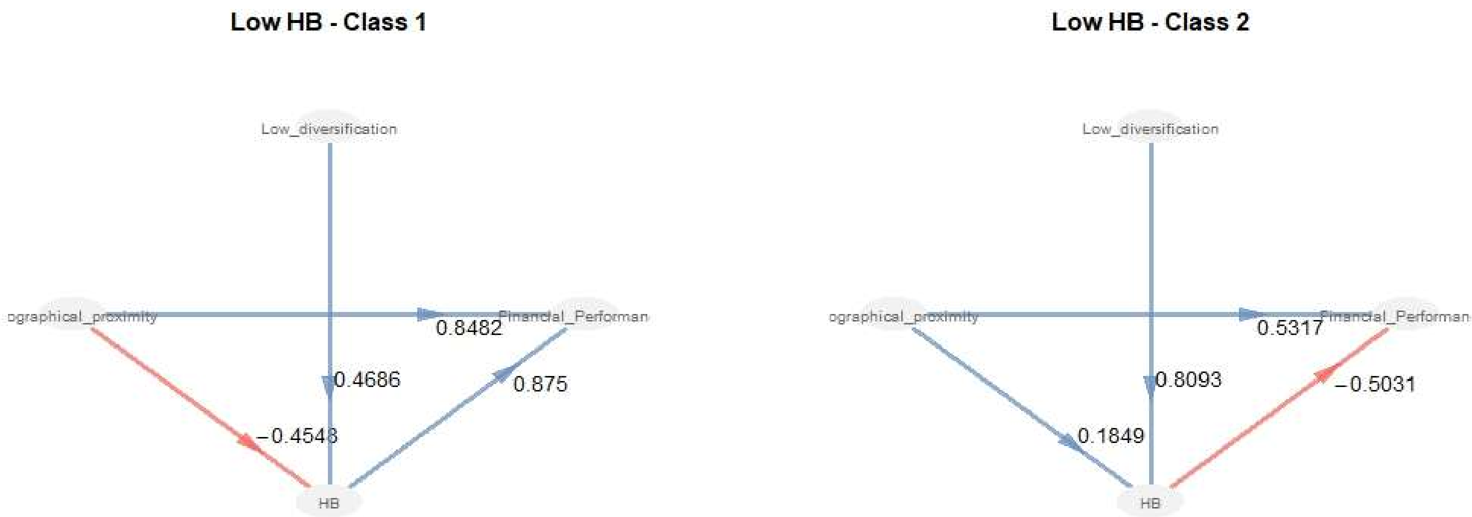

FIGURE B.3: Local Model Low Home Bias: Classes 1 and 3

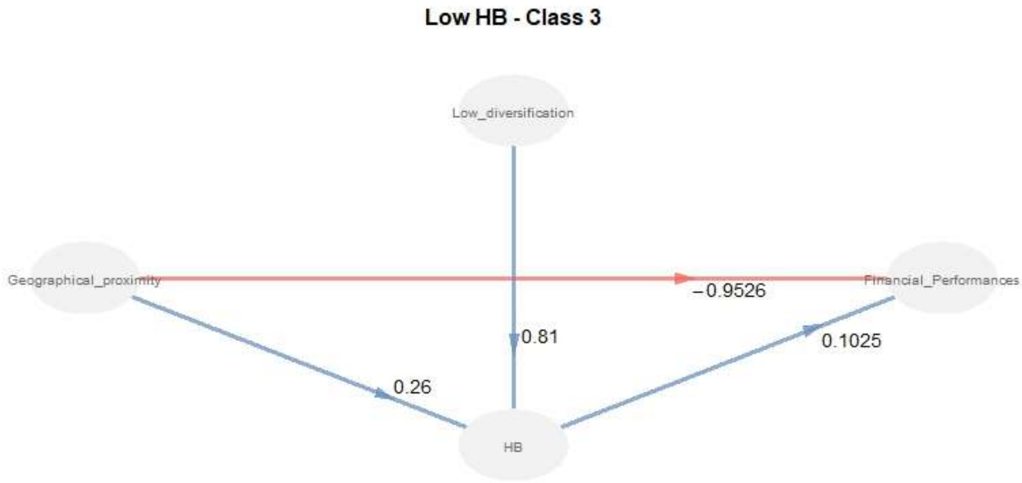

FIGURE B.4: Local Model Low Home Bias: Class 3

B.2 Figures of Chapter 7

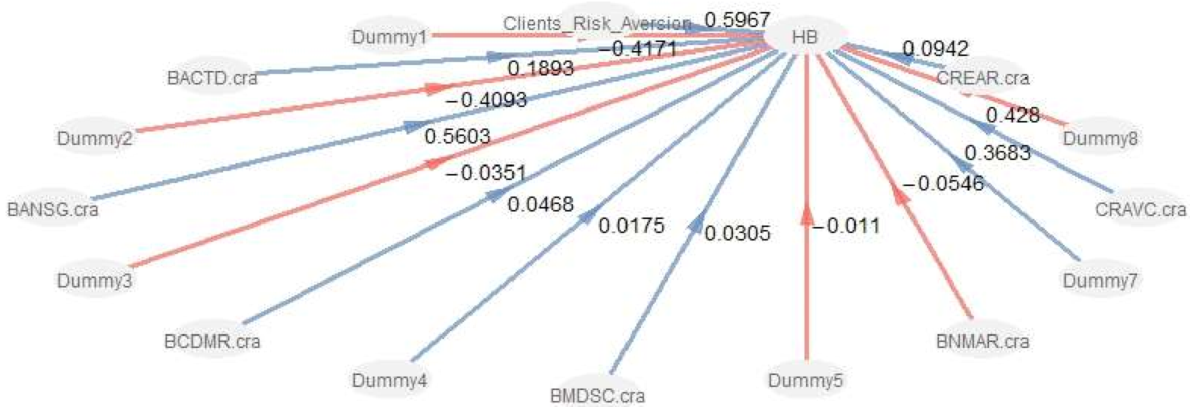

FIGURE B.5: Moderators in the relationship between Geographical Proximity and Home Bias

Bibliography

[1] Elton E. J. “Modern portfolio theory and investment analysis”. // *John Wiley Sons* (2009).

[2] Rigoni U. “Finanza comportamentale e gestione del risparmio”. // *G.Giappichelli Editore - Torino* (2006).

[3] Kumar A. Goetzmann W. “Why Do Individual Investors Hold Under-Diversified Portfolios?” // *Yale School of Management Working Papers* (2005).

[4] Tversky A. Kahneman D. “Prospect theory: An analysis of decision under risk”. // *Econometrica: Journal of the econometric society* (1979), pp. 263–291.

[5] Tversky A. Kahneman D. “Advances in prospect theory: Cumulative representation of uncertainty”. // *Journal of Risk and uncertainty* 5.4 (1992), pp. 297–323.

[6] Poterba J. French K. “Investor Diversification and International Equity Markets”. // *American Economic Review* 81.2 (1991), pp. 222–226.

[7] Nationwide Mutual Insurance Company. “Reducing home bias and portfolio volatility through global investing”. // *White Paper(Nationwide Mutual Insurance Company)* ().

[8] Weber M. Kilka M. “Home bias in international stock return expectations”. // *The Journal of Psychology and Financial Markets* 1.3.4 (2000), pp. 176–192.

[9] Mondria Dziuda W. “Asymmetric Information, Portfolio Managers, and Home Bias”. // *Rev Financ Stud* 25.7 (2012), pp. 2109–2154.

[10] De Moor L. Vanpèe R. *Bond and Equity Home Bias and Foreign Bias: an International Study.* Tech. rep. Hogeschool-Universiteit Brussel, Faculteit Economie en Management, 2012.

[11] Macchiavelli M. De Marco F. *The political origin of home bias: The case of Europe*. Tech. rep. Washington: Board of Governors of the Federal Reserve System, 2016.

[12] Meade E.E. Mann C.L. *Home bias, transactions costs, and prospects for the Euro: A more detailed analysis*. Research Note 6. Deutsche Bank Research, 2002.

[13] Rogoff K. Obstfeld M. "The six major puzzles in international macroeconomics: is there a common cause?" // *NBER Annula Conference on Macroeconomics* (2000), pp. 339–390.

[14] Ananth J. Domowitz I. Madhavan G. "Liquidity, Volatility and Equity Trading Costs across Countries and over Time". // *International Finance* (2001), pp. 55–221.

[15] Miguel A. F. Ferreira M.A. "The determinants of domestic and foreign bond bias". // *Journal of Multinational Financial Management* (2011), pp. 279–300.

[16] Yan I.K. Faruqee H. Li S. *The determinants of international portfolio holdings and home bias*. Tech. rep. Iternational Monetary Fund, 2004.

[17] Boot A.W.A. "Relationship banking: What do we know?" // *Journal of financial intermediation* 9 (2000), pp. 7–25.

[18] Bosch T. Schoenmaker D. "Is the home bias in equities and bonds declining in Europe?" 2008.

[19] *Raccolta bancaria a mezzo di obbligazioni, prospetto e regole di condotta*. CONSOB, 2011.

[20] Tarantola A. "Il Credito Cooperativo di domani: sviluppo, efficienza e solidarietà". // *XIV congresso nazionale del credito cooperativo, Banca d'Italia* (2011).

[21] Hult G. Tomas M. Hair Jr. Joseph F. "PLS-SEM: Indeed a silver bullet". // *Journal of Marketing Theory and Practice* 19.2 (2011), pp. 139–152.

[22] Wold H. "Path modles with latent variables: The NIPALS approach". // *Quantitative sociology* (1975), pp. 307 –357.

[23] Lohmöller J-B. *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg, 1989.

[24] Dongili P. Zago A. "Financial crisis, business model and the technical efficiency of Italian Banche di Credito Cooperativo". // *Department of Economics Sciences - University of Verona* (2014).

[25] Giorgia Simon. "What Drives the Concentration of Household's Investments in Bank Bonds?" // *Department of Management - Ca' Foscari University of Venice* (2015).

[26] Henseler W.W. Wang H. Vinzi E. Chin V. *Handbook of Partial Least Squares : Concepts Methods and Applications*. Springer, 2010.

[27] Sanchez G. *PLS Path Modeling with R*. Ed. by Trowchez Editors. Berkeley, 2013.

[28] Kenny D.A. Baron R.M. "The moderator-mediator variable distinction in social psycological research: conceptual, strategic, and statistical consideration". // *Journal of Personal and Social Psycology* 51.1173-1182 (1986).

[29] Judd C.M. McClelland G.H. "Statistical difficulties of detecting interactions and moderator effects". // *Psycological Bulletin* 114 (1993), pp. 376–390.

[30] Russolillo G. Trinchera L. "On the use of Structural Equation models and PLS Path Modeling to built composite indicator". // *Macerata University* (2010).

[31] Gutiérrez-Nieto B. Cuéllar-Fernandez B. Serrano-Cinca C. Fuertez-Callén Y. "Path modelling to bankruptcy: causes and symptoms of the banking crisis". // *Centre Emile Bernheim* 007.11 (2011).

[32] Wikipedia. *Federazione Veneta BCC*. (Visited on 2017).

[33] Bechger T.M. Hox J.J. "An introduction to Structural Equation Modeling". // *Family Science Review* 11.354 - 373 (1988).

[34] Henseler J. "On the convergence of the partial least squares path modeling algorithm". // *Computational Statistics* 25.1613-9658 (2010), pp. 107–120.

[35] Leisch F. Monecke A. "emPLS: structural equation modeling using partial least squares". // *Journal of Statistical Software* (2012).

[36] Hult G. Tomas M. Hair Jr. Joseph F. *A primer on partial least squares structural equation modeling (PLS-SEM)*. 2016.