UNIVERSITÀ DEGLI STUDI DI PALERMO
Dottorato in Scienze Economiche, Statistiche, Psicologiche e Sociali
Dipartimento di Scienze Economiche, Aziendali e Statistiche
Settore Scientifico Disciplinare: SECS-P/05

Essays on financial stability: an analysis based on NUTS2 and NUTS3 data for Italy

IL DOTTORE
Fabio Parla

IL COORDINATORE
Prof. Vito M.R. Muggeo

IL TUTOR
Prof. Andrea Cipollini

CICLO XXX
ANNO CONSEGUIMENTO TITOLO 2018

Acknowledgements

First, I want to thank my Tutor, Professor Andrea Cipollini, for his guidance and help he has given me over the past three years. His experience and scientific knowledge have been extremely beneficial for my PhD. He always encouraged me to have a critical approach on the research challenges I faced throughout my PhD. I also thank him for the opportunity he has given me to do research experiences abroad, which enable me to grow up not only as a student, but also as a person. This thesis would not have been possible without his extremely important contribution.

I want to thank the Chair of the PhD program, Professor Vito M.R. Muggeo, and the Head of the Department of Economics, Business and Statistics (DSEAS), Professor Marcello Chiodi. Special thanks also to Professor Pietro Busetta who has always encouraged me to follow my interest in research.

Part of this thesis has been written during different visiting research periods I spent at the Essex Business School (EBS, University of Essex) and at the Queen Mary University of London (QMUL). As for the EBS, I would like to thank Professor Claudia Girardone for supervising my research period. I thank Professor Simon Price for his useful comments on a research project. I want to thank also Professor Haroon Mumtaz for his supervision at the QMUL and for the helpful comments he has given me during that period.

Thanks also to Marco, Pietro, Marianna, Ola, Ieva and the other PhD students for the interesting conversations and, in general, the nice time we had together during this three years.

Special thanks to my family for their unconditional support, and to Marta, my love, She has always encouraged me to do my best.

Contents

List of Figures

List of Tables

Introduction

This thesis is a collection of three essays on financial stability for Italy. The focus is on financial distress, proxied by region specific default rates and, on two main determinants of financial distress such as perturbations to credit supply and those to the housing market.
The Italian economy has always been characterized by the divide between Center-North (more economically developed) and Mezzogiorno (less economically developed).[1] Consequently, the analysis of financial stability is based on macroeconomic and financial variables (publicly-available) at both regional and provincial level. According to the classification of Nomenclature of territorial units for statistics (NUTS), the NUTS2 level classifies different groups of regions (NUTS1) into regional entities, while a more disaggregated classification involves grouping different regions into provincial entities (NUTS3) (see EUROSTAT, 2015).
Most of the data used in this thesis are from publicly-available datasets. As for macroeconomic variables, including for example value added and employment to population ratio, I use data from the Statistical database of the Italian National Institute of Statistics, while data on financial market variables, such as information on credit quantities and quality of credit, are from the Statistical database of Bank of Italy. I also use a confidential and unique dataset containing information on house prices (at municipal level) provided by the Real Estate Market Observatory managed by the Italian Revenue Agency ("Agenzia delle Entrate - Osservatorio del Mercato Immobiliare"), combined with publicly-available data on the number of transactions (Italian Revenue Agency).

The recent financial crisis of $2007 - 2008$ has exacerbated the disparities among the different parts of the country. As for the real economy, at the end of 2015 the Mezzogiorno regions have experienced a reduction of 13 point percentage of the GDP (or value added) relative to its pre-crisis period (say 2007), twice the decline recorded in the Center-Northern regions (Bank of Italy, 2015). Furthermore, whilst the level of employment in Center-North has returned to its pre-crisis standard, after the reduction

[1] Mezzogiorno includes the six Southern regions and the Islands of Sardinia and Sicily.

reported during the last seven years, the number of employees in Mezzogiorno regions is still below its pre-crisis level. The macro-regional gap in real economic activity is combined with important differences in terms of access to credit from the banking sector given that the Italian financial system is typically bank-based.

These stylized facts regarding the recent crisis period motivate Chapter 1 which focuses on credit supply (disaggregated only at macro-regional level). Therefore, in Chapter 1, my focus is on the estimation of the effects of macro-regional shocks arising in the Italian credit market (both from demand and supply sides) to the real economic activity in Italy over the $2008-2014$ prolonged crisis period (given the lowest peak reached by the employment to population ratio in the last year of the observed sample). To disentangle credit demand and supply shocks, I use a structural Vector Autoregression (VAR) model fitted to three endogenous variables, including: the loan interest rate, the loans growth rate and the employment to population ratio, whose data are observed at annual frequency, for 103 Italian provinces.
The structural shocks are identified through heteroscedasticity (Rigobon, 2003; Lanne & Lütkepohl, 2008), by letting the variance of the shocks to switch across four Italian macro-regions: North, Centre, South and Islands. Once obtaining the structural shocks, the economic interpretation is achieved by using ex post theory-driven sign restrictions. The empirical findings suggests that, during the $2008-2014$ crisis period, the Northern regions are those facing a milder credit crunch than the ones recorded in the rest of Italy. Moreover the consequences for the real economic activity due to a credit contraction are investigated only for the country as whole.

Housing market dynamics is one of the major driver of financial stability. The narrowest (geographical) focus of this thesis is on the analysis of local housing market resilience to fundamentals shocks. The analysis in Chapter 3 examines how different is the response of house prices and sales volumes in relatively small geographical areas (restricted to a main regional capital and its neighbours) to a one standard deviation negative housing demand shock in a main regional capital.
More specifically, in Chapter 3, I focus on the spatio-temporal diffusion of house prices and transaction volumes spillovers across 93 Italian provinces, over the period $2004-2016$. The aim of the analysis is threefold. First, I investigate the transmission mechanism of house prices spillovers across space and time – known as "ripple effect" – by extending the information set to transaction volumes. Second, the econometric strategy shaped in this chapter enables to assess the heterogeneity in the spatial-temporal diffusion, by exploring how shocks (e.g. negative housing demand shocks) originating in 10 main regional capitals spill over to neighbouring provinces. Finally, I also contribute to the literature on house price-volume correlation, by observing their co-movements in

response to an unobserved shock to housing demand.
For this purposes, I use the above mentioned dataset on house prices, combined with publicly-available data on the number of transactions (Italian Revenue Agency). I use a Global VAR (GVAR) model, fitted to house prices and sales (as endogenous variables) together with a spatial exogenous regressor, where the structural housing demand shock is identified by using sign restrictions (Eickmeier & Ng, 2015).
The results provide evidence of a strong "ripple effect" in transaction volumes for almost all the 10 observed Italian provinces, while the house prices spillovers across regional capitals and neighbours display low magnitude, with the only exception of Roma. The empirical findings reveal housing markets segmentation as suggested by the different response of neighbours house prices and sales volumes to a one standard deviation negative housing demand shock in a main regional capital. Moreover, sales volumes are more sensitive than house prices to a negative housing demand shock.

While the analysis of the propagation of housing demand shock or the one for credit supply innovation has a structural form modelling flavour, the investigation (see Chapter 2) of financial distress spillover across regions has a reduced form flavour. The analysis is motivated by the strong consolidation process characterizing the Italian banking system since 1990's. One of its main consequences has been the acquisition of troubled banks located in Mezzogiorno by Northern banks, with the subsequent loose of autonomy in Mezzogiorno banks (Papi *et al.*, 2015). However, the operations of mergers and acquisitions (M&As) have not completely reduced the exiting gap in the conditions of the macro-regional financial systems. There are still not-negligible differences, with the Mezzogiorno regions reporting a worse quality of credit and a higher cost of credit than the other parts of the country (Bank of Italy, 2017).
More specifically, in Chapter 2, the empirical analysis focuses on the presence of spatial spillovers across Italian regions, using the default rates on loans facilities as proxy of the loans probability of default, for three private sub-sectors: consumer households, non-financial firms and producer households.[2] The quarterly series on loan default rates cover the $1996-2015$ time span. In particular, the aim of this chapter is twofold. First, I investigate the presence of spatial dependence across the regional loan default rates. Second, I evaluate whether the Mezzogiorno regions are more affected by spillover effects arising from Northern regions. For this purpose, I use the connectedness measures proposed by Diebold & Yilmaz (2012, 2014) and, more recently, by Greenwood-Nimmo *et al.* (2015). These approaches rely on the construction of the Generalized Forecast Error Variance Decomposition (GFEVD) obtained through the estimation of VAR models

[2]According to the definition provided by Bank of Italy, producer households are defined as individual firms, informal partnership and unregistered company, producers of marketable goods and financial services with up to five employees; activities auxiliary to financial intermediation without employees.

(one for each of the three private sub-categories), fitted to the 20 regional default rates series. Since the relative large number of endogenous variables, I use the Adaptive Elastic net shrinkage estimator.
As for the first issue, the empirical findings reveal an increase in default rates spatial dependence over the $2011Q4 - 2015Q4$ (crisis) period, especially for producer households. As for the second issue, I find evidence of a strong dependence of the Islands (two Mezzogiorno regions) from the North of Italy, while the other Southern regions are found to be the most contributor, together with the Northwest of Italy, of financial distress to the remaining macro-regions.
The major findings are evidence of spillover effects during the most recent crisis period $(2011-2015)$ and, in particular, the major vulnerability of the Northeast (together with Insular Italy) to financial distress shocks.

Bibliography

Bank of Italy. 2015. *The economy of the Italian regions.* Regional Economies No. 43. Bank of Italy.

Bank of Italy. 2017. *The economy of the Italian regions.* Regional Economies No. 23. Bank of Italy.

Diebold, Francis X, & Yilmaz, Kamil. 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, **28**(1), 57–66.

Diebold, Francis X, & Yilmaz, Kamil. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, **182**(1), 119–134.

Eickmeier, Sandra, & Ng, Tim. 2015. How do US credit supply shocks propagate internationally? A GVAR approach. *European Economic Review*, **74**, 128–145.

EUROSTAT. 2015. *Regions in the European Union.* Manuals and guidelines. European Commision, Available at `http://ec.europa.eu/eurostat/web/nuts`.

Greenwood-Nimmo, Matthew, Nguyen, Viet Hoang, & Shin, Yongcheol. 2015. *Measuring the connectedness of the global economy.* Working paper No. 7/15. The University of Melbourne.

Lanne, Markku, & Lütkepohl, Helmut. 2008. Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, **40**(6), 1131–1149.

Papi, Luca, Sarno, Emma, & Zazzaro, Alberto. 2015. *The geographical network of bank organizations: issues and evidence for Italy.* Working paper No. 105. Mo.Fi.R.

Rigobon, Roberto. 2003. Identification through heteroskedasticity. *Review of Economics and Statistics*, **85**(4), 777–792.

Chapter 1

Credit Demand and Supply Shocks in Italy during the Great Recession

1.1 Introduction

A predominant feature of the Great Recession has been a prolonged contraction of credit to the private sector in a number of countries. The aim of this chapter, based on provincial level data, is twofold. First, I focus on the identification of credit demand and supply shocks in explaining the credit contraction in Italy. Second, I am also interested in analyzing the effects of the identified credit supply shock on real activity for the Italian economy.

The slowdown in bank lending which occurred in many advanced economies has led to a debate about the effects of disturbances in credit markets on business cycles. In spite of the increasing importance of capital markets, the Euro financial system is typically bank-based. Furthermore, bank loans play a non-negligible role in the financing of private investment and consumption in the European countries. The Italian financial system has been dominated by banks: the ratio of total loans to the Italian banks total assets was 57.6 percent at the end of 2014.[1] Hence, bank lending might play an important role in explaining fluctuations of economic cycle. In the aftermath of the financial crisis, the Italian banking system has seen a slackening growth of bank loans to non-financial corporations and households. The Italian year-on-year growth rate of loans to private sector fell from 9.2 percent in the first quarter of 2008 to 1.1 percent in the first quarter

[1] Data on balance sheet of banks resident in Italy are collected from the Statistical database of Bank of Italy.

of 2010. After a sharp upturn, the growth rate has become negative since the second quarter of 2012.

A number of empirical studies based on micro-data informing on bank-firm relationship employ the methodology proposed by Khwaja & Mian (2008) to identify credit supply shocks (see Bonaccorsi di Patti & Sette, 2016, for the Italian economy, among the others). The Khwaja & Mian (2008) methodology exploits a sample which includes observations for a pre and post crisis period, and it is based on the estimation of a regression of the change in the loans provided by each bank to its borrowing firms after an exogenous shock (e.g. a crisis event) as a function of bank exposure to that shock. Del Giovane *et al.* (2017) use Bank Lending Survey (BLS) to identify, through zero exclusion restrictions, the simultaneous equations system fitted to interest rates and loans data for Italy.

While the previous studies are only interested in the identification of a credit supply factor, some authors are also concerned with their real effect using a two-stage estimation analysis. Cingano *et al.* (2016) use data on bank-firm relationship and they identify credit supply shocks through the variation in bank reliance on the interbank market at the end of 2006, leading to different bank exposure to the July 2007 liquidity shock. The proxy used by Cingano *et al.* (2016) for the real activity is the private investment. The study of Barone *et al.* (2016) uses bank-province relationship data for the Italian economy to identify a local (province) credit supply indicator. In a second stage of the analysis, they assess the impact of credit supply shock on investment, value added and employment.

The Khwaja & Mian (2008) methodology employed by Cingano *et al.* (2016) and by Barone *et al.* (2016) relies on an individual bank exposure to an exogenous shock (e.g. crisis event) switching from a no crisis period to one characterized by turmoil.

Since I focus only on a prolonged crisis time span, I exploit the heterogeneity in the data across Italian macro-regions. More specifically, ex-ante, I employ identification through heteroscedasticity (see Rigobon, 2003; Lanne & Lütkepohl, 2008) and, ex-post, I give an economic interpretation to the shocks through sign restrictions (see Mumtaz *et al.*, 2015, for a review). In particular, I follow the suggestion of Kick (2016) in setting the sign restrictions: a credit supply (demand) shock moves the price and quantity of credit in opposite (same) directions.

Both methods are popular for the identification of a Structural VAR, which is the model used in this chapter. Moreover, I argue that, contrary to the previous studies which rely on a two-stage analysis, my study, based on an estimation in one-shoot, does not suffer from a measurement error affecting the use of an estimated regressor in the second stage regression.

The empirical findings show that credit supply shocks play a more important role than innovations to demand for credit. Furthermore, there is evidence that credit crunch hits the North of Italy less than the remaining macro-regions, especially the South-Italy. The chapter is structured as follows. Section 1.2 provides a literature review on identification of shocks to credit markets. Section 1.3 describes the empirical methodology. Section 1.4 describes data and the empirical findings and Section 1.5 concludes.

1.2 Literature review

Credit market shocks can be distinguished in credit supply and credit demand shocks. Credit supply shocks are exogenous innovations to credit supply which affect the capability or willingness of banks to lend. They might reflect changes in regulatory capital requirements or in risk aversion of banks and other financial institutions. Moreover, the increasing diffusion of financial market innovations such as instruments of credit risk transfer that could affect banks willingness to lend has been labelled as credit supply shocks (Atta-Mensah & Dib, 2008; Halvorsen & Jacobsen, 2014). Furthermore, other factors may affect the supply of credit. For example, these factors include the availability and price of banks funding and the competition degree of banking sector (ECB, 2011). Oppositely, credit demand shocks are exogenous innovations to credit demand mainly depending on the macroeconomic outlook. Consequently, changes in macroeconomic equilibrium influence the balance sheet of non-financial corporations and households, affecting borrowers preferences in the volume of bank credit demanded (Peersman, 2011; Kick, 2016).

As mentioned in the introduction, a number of empirical studies on the Italian credit crunch are only interested in the identification of credit supply shocks. Presbitero *et al.* (2014) relies on the identification of constrained Italian firms, using firms survey data containing information on loan applications and bank decisions. The authors main focus is the role played by functional distance between the loan office and the headquarters where final lending decisions are made to explain the tightening in lending conditions in Italy. For this purpose, the authors combine survey data on firms with aggregate data on banks informing on the openings and closures of branches at the bank-province level. The authors use a sample of monthly observations from 2008:1 to 2009:4 and the empirical findings show that the credit crunch experienced in Italy after Lehman Brothers collapse has been more severe in provinces with larger shares of branches owned by distantly managed banks. Moreover, there is evidence of a home bias, given that the credit crunch has not been harsher for small and economically weak firms.

The identification methodology put forward by Khwaja & Mian (2008), which is based on Credit Register data for firms that have multiple lenders, has been applied by a number of studies. In particular, the study of Khwaja & Mian (2008) seeks to evaluate the effects of bank liquidity shocks due to the nuclear test of Pakistan in 1998 on its economy by separately estimating the bank lending channel, that is the inability of banks to cushion borrowing firms against bank-specific liquidity shocks, and the firm borrowing channel, that is the inability of firms to smooth out bank lending channel effects by borrowing from alternative sources of financing.

The approach proposed by Khwaja & Mian (2008) consists of estimating a regression of the change in the loans provided by each bank to its borrowing firms after an exogenous shock (e.g. the nuclear test) as a function of bank exposure to that shock. For this purpose, they use firm fixed effects to capture shifts in the demand for loans and other unobservable borrower characteristics, such as changes in their balance sheet conditions. The identification methodology provides an estimate of the differential change in credit supply for the same firm, associated with a different exposure of the lending banks to the exogenous shock. Albertazzi & Marchetti (2010) present evidence of a contraction of credit supply associated to low bank capitalization and scarce liquidity, over the 6-month period following the Lehman bankruptcy. Bofondi *et al.* (2013) exploit the differential exposure to the sovereign risk between domestic banks and foreign banks operating in Italy. The authors find that the lending of domestic banks grew less (and their interest rates were higher) than that of foreign banks, after the outbreak of the sovereign debt crisis. Bonaccorsi di Patti & Sette (2016) link banks' balance sheet conditions to the provision of credit and show that Italian banks that relied heavily on securitization prior to the subprime crisis curtailed lending more than other banks.

Del Giovane *et al.* (2017) estimate a system of two simultaneous equations regarding the interest rates and loan amounts of 11 Italian banks. The authors use demand and supply dummies obtained from the Eurosystem Bank Lending Survey (BLS).[2] In order to identify the simultaneous equations system, the demand dummies are excluded from the equation where the dependent variable is the price and the supply factor dummies are excluded from the equation involving quantity. After a number of robustness checks, the authors acknowledge that they cannot exclude the possibility that their findings are affected to some extent by some residual endogeneity. The authors find that the effects of the supply restriction on both the cost and the availability of credit were, on average, stronger during the sovereign debt crisis than during the Lehman global crisis. Moreover, the authors find that credit crunch was mostly related to the banks' risk perception during the global crisis, whereas funding conditions became predominant during

[2]Ciccarelli *et al.* (2015) use the Bank Lending Survey (BLS) for the Euro area and the Senior Loan Officer Survey (SLOS) for the U.S. Contrary to the study of Del Giovane *et al.* (2017) which employs BLS data for Italy only to identify credit demand and credit supply shocks, Ciccarelli *et al.* (2015) are also interested in the real effects of credit supply shocks. The qualitative data are transformed into quantitative and treated as endogenous variables together with proxies of output, prices and monetary policy rates in a Vector Autoregression model, VAR, fitted to the Euro area and to the US separately.

the sovereign debt crisis.

The second empirical issue regarding the impact of the identified credit supply shock on real activity of the Italian economy has been addressed by the following studies. The study of Cingano *et al.* (2016) uses in the first stage the Khwaja & Mian (2008) identification methodology. In particular, the authors use data on bank-firms relationships and they identify credit supply shocks through the variation in bank reliance on the interbank market at the end of 2006, leading to different bank exposure to the July 2007 liquidity shock. The authors' findings show that, although credit tightening was homogeneous across firms, investment fell by a much larger amount among smaller and younger firms, and those with higher bank dependence. Bottero *et al.* (2015) show that the Greek bailout in 2010 led to a fall in loan supply in Italy, which depressed investment and employment for smaller Italian firms.

The methodology suggested by Greenstone *et al.* (2014) is used to identify and assess the real effects of a credit supply indicator by Barone *et al.* (2016). The authors use confidential data over 2008-2011, obtained from the Bank of Italy Supervisory Report, on total outstanding loans extended by Italian banks to the private sector (firms and households) aggregated into local credit markets corresponding to provinces. The identification strategy employed by Barone *et al.* (2016) is based on data capturing bank-provinces relationships (hence it is similar to Khwaja & Mian, 2008). More specifically, the authors focus on the identification of a local (province) credit supply indicator by, first, using a panel regression. The dependent variable is the change in credit granted by one of the 650 banks to households and firms located in a given province and operating in a given economic sector and the explanatory variables are two dummies. The first dummy measures province-year fixed effects that capture the variation in the change of lending due to local economic factors (capturing local demand). The second dummy measures bank-year fixed effects which identify nationwide bank lending policies. The authors, then, use the coefficient associated to the second dummy and the pre-crisis bank market shares in the province (as weights) to aggregate and to construct a province-year credit supply index. In a second stage, the credit supply real effects are estimated regressing either value added, or investment, or employment (observed for each province) on the estimated local credit supply variable. The empirical findings show that the most severe effect of the credit crunch occurred in the North and Central Italy which have firms relatively more dependent on external finance. The methodology suggested by Greenstone *et al.* (2014) is also employed by Berton *et al.* (2017), using a matched data set of job contracts, firms and banks in one Italian region (Veneto). The authors, first, identify and construct a credit supply factor at firm level, and, in a second stage, they assess the impact on employment. The empirical findings (for Veneto region) show that the effects of the credit crunch have been particularly severe for smaller, younger and less productive firms, and those with higher debt overhang and weaker bank-firms

relationships have been more vulnerable to the (negative) impact of the credit crunch. Dörr *et al.* (2017) use information on loans by individual banks to firms that borrow from multiple Italian banks, which are exposed to foreign borrowers in distressed countries (Greece, Ireland, Portugal and Spain). The authors use a novel identification method suggested by Amiti & Weinstein (2017) which does not rely on a comparison between access to credit during pre-crisis and a crisis period (as in Khwaja & Mian, 2008), but only on loan data over the 2010-2012 period characterized by Euro sovereign debt crisis. The credit supply and demand components are recovered by imposing an additional constraint. The adding-up constraint states that changes in individual loan growth between banks and firms must add up to the overall, economy-wide change in loan growth. After establishing that credit supply shocks reduce firms' loan growth, Dörr *et al.* (2017) show that credit supply rationing had significant real effects on firms' investment and employment decisions, as well as on total factor productivity. Italian firms with higher exposure to troubled banks reduced their investment and employment and they experienced a significant fall in productivity.
Recent empirical studies on the Italian economy (together with Euro area countries) employ macro-time series data and they identify credit supply shocks and their impact on the real economy by imposing sign restrictions to identify a Structural Vector Autoregression model, SVAR. In particular, Bijsterbosch & Falagiarda (2014) use time-varying parameter Vector autoregression model with stochastic volatility, producing results for Euro area countries, including Italy. The studies of Hristov *et al.* (2012), based on Panel VAR, and the study of Kick (2016), based on Global VAR, analyze the dynamic effect of credit supply on real economic activity in Italy as well as a number of Euro area countries.

1.3 Structural VAR

In this section, I first describe the identification through heteroscedasticity methodology. The first study of identification of structural shocks via changes in volatility is due to Rigobon (2003). Recently, the studies of Lanne & Lütkepohl (2008), Lütkepohl (2012) and Lütkepohl & Netsunajev (2015) show that heteroscedasticity in residuals provides over-identifying restrictions (which can be tested) to traditional SVAR models employed to study the effect of monetary policy shocks. Lütkepohl (2012) identifies shocks by considering changes in volatility in given time periods (with breakpoints specified exogenously). The author considers also a vector generalized autoregressive conditional heteroscedasticity (MGARCH) to model for changes in volatility of residuals. Finally, a third specification model examines changes in volatility by using a Markov regime switching process. Lütkepohl & Netsunajev (2015) use a SVAR to estimate the interaction between US monetary policy and stock market where the identification is obtained by modelling heteroscedasticity in a way similar to Lütkepohl (2012), considering also smooth transition in the variances.

Following Lütkepohl (2005), I carry out with a SVAR analysis, estimating a structural B-model VAR(1) for pooled data, which has the following reduced form representation:

$$y_t = \delta + A_1 y_{t-1} + u_t \tag{1.1}$$

where $y = (interest\ rate_{i,t}, \Delta loans_{i,t}, empl.ratio_{i,t})$ is a $K = 3$ dimensional vector of endogenous variables in province i at time t, namely interest rate on loans (*interest rate*), a log transformation of loans first order difference ($\Delta loans$) and the employment to population ratio (*empl.ratio*), δ is a $K \times 1$ vector of constant terms, A_1 is a $K \times K$ parameter matrix and $u_t \sim N(0, \Sigma_u)$ is a K-dimensional vector of residuals with a non singular covariance matrix $E(u_t u_t') = \Sigma_u$, which is not assumed to be diagonal.

According to Lütkepohl (2005), the B-model specification allows to identify the structural disturbances, ε_t, directly from the VAR residuals. In fact, the white-noise reduced residuals can be expressed as a linear function of the structural disturbances:

$$u_t = B\varepsilon_t \tag{1.2}$$

where B is a non-singular $K \times K$ matrix including the contemporaneous interactions between the endogenous variables and $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$ is a vector of uncorrelated structural shocks. The relationship between the residuals covariance matrix in reduced form, Σ_u, and the structural one, Σ_ε, is:

$$\Sigma_u = E(u_t u_t') = B \underbrace{E(\varepsilon_t \varepsilon_t')}_{\Sigma_\varepsilon} B' \tag{1.3}$$

hence, the structural form of VAR is:

$$y_t = \delta + A_1 y_{t-1} + B\varepsilon_t \tag{1.4}$$

Since the structural shocks are mutually uncorrelated, it is possible to achieve an economic interpretation of these shocks affecting the endogenous variables of the VAR. However, the identification of the structural parameters of the model, B and Σ_ε, involves additional identifying assumptions based on institutional knowledge, economic

theory, or other extraneous constraints on the model responses (see Kilian, 2013).
Let me consider the relationship between the reduced and structural forms of the residuals covariance matrix in eq.(1.3):

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ & s_{22} & s_{23} \\ & & s_{33} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & 0 & 0 \\ 0 & \sigma_{22}^2 & 0 \\ 0 & 0 & \sigma_{33}^2 \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \tag{1.5}$$

Since the reduced covariance matrix is symmetric, the system in eq.(1.5) has $K(K+1)/2$ equations and $K(K+1)$ unknown parameters.
This identification issue has been faced by authors by proposing different approach (see Kilian, 2013, for a review of the main identification approaches employed in the SVAR framework).
To identify the structural parameters of the model, I follow the methodology suggested in the studies of Lanne & Lütkepohl (2008) and Ehrmann *et al.* (2011) which relies on the identification approach originally proposed by Rigobon (2003).
In particular, I estimate a reduced-form of a VAR(1) by Ordinary Least Squares (OLS) for each equations separately. Once pooling the data in a K-dimensional vector of time series, the eq.(1.1) can be written compactly as follows:

$$Y = \Theta Z + U \tag{1.6}$$

where $Y = (y_1, \ldots, y_T)$ is a $K \times T$ vector of endogenous variables, $\Theta = (\delta, A_1)$ is a $K \times (K+1)$ matrix of coefficients, including the intercepts, $Z = (Z_0, \ldots, Z_{T-1})$ is a $(K+1) \times T$ matrix of lagged variables, with $Z_t = (\boldsymbol{1}, y_t)$, and $U = (u_1, \ldots, u_T)$ is a $K \times T$ vector of residuals (see Lütkepohl, 2005, for further details).
Alternatively, eq.(1.6) can be written as:

$$\begin{aligned} \text{vec}(Y) &= \text{vec}(\Theta Z) + \text{vec}(U) \\ y &= (Z' \otimes I_K)\ \theta + u \end{aligned} \tag{1.7}$$

where $y = \text{vec}(Y)$ and $u = \text{vec}(U)$ are the $(KT \times 1)$ vectors of variables and residuals, respectively, and $\theta = \text{vec}(\Theta)$ is a $(K^2 + K) \times 1$ vector of VAR coefficients.
The multivariate OLS estimator $\hat{\theta}$ is computed as follows:

$$\hat{\theta} = ((ZZ')^{-1}Z \otimes I_K)y \tag{1.8}$$

The estimation of eq.(1.8) is equivalent to the OLS estimation applied equation by equation, separately (Lütkepohl, 2005).
To retrieve the structural representation of the VAR model, the one in eq.(1.4), I need to establish different regimes of volatility. This allows the determination of the covariance matrix structures as well as identifying the system of equations.
Regimes of volatility are selected on the basis of geographical discrimination. Particularly, four heteroscedastic regimes are defined, corresponding to different Italian macro-areas: North Italy, Central Italy, South Italy and Insular Italy.
The sample of observations is divided into 4 sub-samples, based on geographical characteristics, $S = (S_{North\ Italy}, S_{Central\ Italy}, S_{South\ Italy}, S_{Insular\ Italy})$.
Constructing the covariance matrix structures is carried out by choosing the North Italy as the first regime, whereas the other regimes are: *(i)* Central Italy, *(ii)* Southern Italy and *(iii)* Insular Italy.[3]
The covariance matrix structure has the following representation:

$$\Sigma_1 = BB', \quad \Sigma_i = B\lambda_i B', \qquad i = 2, \ldots, 4 \tag{1.9}$$

where

$$\Sigma_1 \quad \text{for } i \in S_{North\ Italy} \quad \text{and} \quad \Sigma_i = \begin{cases} \Sigma_2 & \text{for } i \in S_{Central\ Italy} \\ \vdots & \\ \Sigma_4 & \text{for } i \in S_{Insular\ Italy} \end{cases} \tag{1.10}$$

Once the reduced form of VAR(1) model is estimated by OLS estimation, the corresponding residuals are used in order to estimate the unknown parameters.
The set of unknown parameters includes matrix B coefficients and the variances of the structural error terms.
Assuming normality of the error terms, the structural parameters are obtained by Maximum Likelihood (ML) estimation. The Multivariate Gaussian log-density function at time t and for macro-region i is:

[3] In different exercises, I use various combinations in defining the heteroscedastic regimes by setting other macro-regions as first regime. The results of this exercises confirm the ones I show in the rest of the chapter.

Table 1.1: Theory-driven ex post sign restrictions on B matrix

Impact on	*Credit demand shock*	*Credit supply shock*	*Real shock*
interest rate	-	+	***n.a***
$\Delta loans$	-	-	***n.a***
empl. ratio	-	-	-

Note. Here the sign restrictions are related to negative shocks

$$log\ l(B,\lambda) = -\frac{KT}{2}\ log(2\pi) - \frac{1}{2}\sum_{i=1}^{4}|log(\Sigma_i)| - \frac{1}{2}\sum_{i=1}^{4}(u_i'\Sigma_i^{-1}u_i) \tag{1.11}$$

where Σ_i is the covariance matrix of the reduced-form residuals, expressed in terms of the structural form coefficients as described in eqs.(1.9) and (1.10).[4] As mentioned above, identification through heteroscedasticity is only a statistical tool, and to give an economic interpretation of the structural form shocks I use, ex post, sign restrictions on each column of the impact multiplier matrix B (see Table 1.1)

The economic identification of credit demand and supply shocks is based on a minimal set of identifying restrictions in line with previous studies (Peersman, 2011; Barnett & Thomas, 2013; Kick, 2016). More specifically, a negative credit demand shock reduces both credit price and the amount of loans. Conversely, a negative credit supply shock produces an increase of the loan interest rate as well as reducing the quantity of bank lending (see Hristov *et al.*, 2012). The real variable is affected by credit supply and demand shocks negatively. Following Kick (2016), I do not expect any prior sign restriction from the responses of the credit variables to the real shocks.

Since the number of unknowns is equal to eighteen and the number of moment conditions (see eqs.(1.9) and (1.10)) is equal to twenty-four equations, a Likelihood Ratio test is employed to test for the six over-identifying restrictions. Once defining $\hat{\theta} = (\hat{B}, \hat{\lambda}_i)$, for $i = 1, \ldots, 4$, as the ML estimation of the structural parameters, I compute the LR test as follows:

$$LR = -2[ln(\hat{\theta}_R) - ln(\hat{\theta}_{UR})] \tag{1.12}$$

where $\hat{\theta}_R$ is the ML estimator of the restricted model and $\hat{\theta}_{UR}$ is the ML estimator of the unrestricted model. Under the null hypothesis, the Likelihood ratio statistic has an asymptotic χ^2 distribution with degree of freedom equal to the number of the over-identifying restrictions.

[4]The log density functions are generated by using the **mvtnorm** package in R. The optimization problem is solved by minimizing the negative of the sum of the log densities by using the “BFGS” method. The “BFGS” method is a quasi-Newton method which uses function values and gradients to build up a picture of the surface to be optimize.

I also carry out with a structural dynamic analysis.
Let me consider the reduced form of VAR(1) model in eq.(1.1) and rewrite the equation as follows:

$$A(L)y_t = \delta + u_t \tag{1.13}$$

where $A(L) = [I_K - A_1 L]$ is the corresponding autoregressive lag order polynomial. In case of a stationary process, the matrix polynomial in the lag operator can be inverted, $[I_K - A_1 L]^{-1}$, and the VAR(1) admits a Vector Moving Average representation of an infinite order, VMA(∞):

$$\begin{aligned} y_t &= [I_K - A_1 L]^{-1}\delta + [I_K - A_1 L]^{-1}u_t \\ y_t &= (I_K + \Psi_1 + \Psi_2 + \ldots)\delta + (I_K + \Psi_1 L + \Psi_2 L^2 + \ldots)u_t \\ y_t &= (I_K + \Psi_1 + \Psi_2 + \ldots)\delta + (\underbrace{B}_{\Phi_0} + \underbrace{\Psi_1 B}_{\Phi_1} L + \underbrace{\Psi_2 B}_{\Phi_2} L^2 + \ldots)\varepsilon_t \end{aligned} \tag{1.14}$$

where $u_t = B\varepsilon_t$, $\Psi_i = A_1^i$ (with $\Psi_0 = I_K$) are the coefficient matrices of the reduced VMA(∞) representation, and $\Phi_i = A_1^i B$ are the coefficient matrices of the structural VMA(∞) representation. From the structural VMA(∞) representation it is possible to derive the orthogonalized impulse response functions (IRF), which capture the effect of a structural disturbance in the j-th element of ε_t to the i-th variable in y_t after h periods (y_{t+h}):

$$IRF_{h,ij} = \frac{\delta y_{i,t+h}}{\delta \varepsilon_{j,t}} = [A_1^h B]_{ij} \tag{1.15}$$

Since $A_1^h B = \Phi_h$, the structural VMA coefficient matrices reflect the impulse responses of the system. The accumulated effect of a structural shock over time is computed from the MA coefficient matrices as the cumulative sum of the corresponding orthogonalized impulse response after h periods.
In my analysis, I compute the cumulative standardized impact of each structural shock over a two-year horizon by estimating $B + A_1 B$. While the standard errors of the parameters of the standardized impact multiplier B are retrieved from the inversion of the Hessian of the maximized log-likelihood function, the confidence intervals for the cumulative impulse response are generated through bootstrap. In particular, for each

regime, I resample 1000 times the estimated residuals of the VAR(1).[5] For each draw, I estimate the parameters of the structural form model by maximizing the log likelihood function.

Another common econometric exercise in SVAR analysis concerns the construction of the historical decomposition of the observed time series. The use of this analysis allows to decompose each endogenous variable into the contribution of its deterministic component (the unconditional mean) and the structural shocks occurring to the K endogenous variables in the system. In particular, for each of 103 Italian provinces, each time observation for the $K = 3$ endogenous variables can be written as follows:

$$\begin{aligned}
&y_2 = \delta + A_1 y_1 + B\varepsilon_2 \\
&y_3 = \delta + A_1 y_2 + B\varepsilon_3 = \delta + A_1(\delta + A_1 y_1 + B\varepsilon_2) + B\varepsilon_3 = \delta + A_1\delta + A_1^2 y_1 + A_1 B\varepsilon_2 + B\varepsilon_3 \\
&\ldots = \ldots \\
&y_T = (\delta + A_1\delta + \ldots + A_1^{T-2}\delta) + A_1^{T-1} y_1 + (A_1^{T-2} B\varepsilon_2 + \ldots + A_1 B\varepsilon_{T-1} + B\varepsilon_T)
\end{aligned} \tag{1.16}$$

which can be written more generally as:

$$y_t = \delta \sum_{j=0}^{t-2} A_1^j + A_1^{t-1} y_1 + \sum_{j=0}^{t-2} A_1^j B\varepsilon_{t-j} \quad , \qquad \text{for } t > 1 \tag{1.17}$$

where $y_1 = y_{2008}$ in my analysis. Constructing the historical decomposition allows to compute the anticipated and unanticipated components of each series.

1.4 Empirical analysis

1.4.1 Data

I use a panel data set of observations which contains information on credit aggregates and a real variable for 103 Italian provinces.

For the purpose of disentangling credit supply shock from the demand-side one, I consider two credit market aggregates and one real activity variable. Hence, as endogenous variables, I use as proxies of price and quantity of credit the loan interest rate and the amount of loans, respectively; the employment to population ratio is the proxy of real

[5] I keep only the replications (which are 421) in line with the ex post identification of the shocks according to the point estimation results.

economic activity. The data are at annual frequency, from 2008 to 2014, for each of 103 provinces, for a total of 2163 observations. I use low-frequency data because of the availability of the employment to population ratio: for each province, data are only made accessible with annual frequency. The shortness of the sample period used is due to the loan interest rate series which starts from 2008.

Information on credit aggregates are from the Statistical Database of Bank of Italy. As for the price of credit, I use the lending rates on loans facilities (stock) series for non-MFI resident sectors. Particularly, I consider the interest rate charged by banks at the end of the fourth-quarter as annual observation.

As for the quantity of credit, I consider the first-order difference of loans to non-MFI resident sectors as endogenous variable.[6] In an attempt to include in my model annual observations instead of quarterly data, I consider the value of loans registered at the end of each fourth-quarter. Taking into account the first difference allows to avoid stationarity problems.

The real aggregate is the employment rate which is defined as the ratio between employed people (aged 15-64) and the corresponding overall resident population. The data are collected from the statistical database of the Italian National Institute of Statistics (ISTAT).

Actually, ISTAT makes available gross value added data which might be used as a proxy for real economic activity at provincial level. Nonetheless, the value added series is not available for 2014.

Since I seek to identify credit supply and demand shocks, and a real shock, through cross sectional heteroscedasticity, I consider four macro-regions: North Italy, Central Italy, South Italy and Insular Italy.[7]

Figures 1.1, 1.2 and 1.3 show the boxplots series of the three endogenous variables for each Italian macro-area from 2008 to 2014. The boxplots provide information on each province which belongs to different macro-regions.

Focussing on the mean values of Figure 1.1, all the Italian macro-regions exhibit the same pattern in the loans interest rate. After a twofold decrease over the 2008-2010, the loan interest rates stabilize around values ranging from 2.8 and 3.5 percent, before exhibiting a temporary upturn in 2011. Afterward, the interest rate on loans values do not exceed 3.7 percent in the 2012-2014 period.

Figure 1.2 shows a more heterogeneous evolution over time of the loan growth rates by

[6] According to the definitions provided by the Bank of Italy, the loans aggregate is defined as the loans disbursed by banks to non-bank sectors. This variable includes mortgage loans, current account overdrafts, loans secured by pledge of salaries, credit card advances, discounting of annuities, personal loans, leasing, factoring, other financial investment (e.g. commercial paper, bill portfolio, pledge loans, loans granted from funds administered for third parties), bad debts and unpaid and protested own bills.

[7] In this analysis, the twenty Italian regions (NUTS2) are grouped into four macro-regions: North Italy (Aosta Valley, Emila-Romagna, Friuli Venezia-Giulia, Liguria, Lombardy, Piedmont, Trentino Alto-Adige and Veneto), Central Italy (Lazio, Marche, Tuscany and Umbria), South Italy (Abruzzo, Apulia, Basilicata, Calabria, Campania and Molise) and Insular Italy (Sardinia and Sicily).

inspection of the boxplots for the macro-regions. Whilst the highest loan growth rate is in South and Insular Italy at the beginning of the crisis, these regions experience the strongest slowdown in the growth rates, starting from 2011. During the last two years of the sample, there is a clear evidence of a recovery in the loans growth rates.
In Figure 1.3, I can observe that during the period 2008-2014, all the four macro-areas exhibit a relevant decline of the employment to population ratio, with different levels of decrease in the territorial areas. Whilst North and Central Italy experience a moderate reduction in the employment to population ratio until 2012 and a moderate upturn in 2013-2014, the South and Insular Italy manifest a significant negative trend during the whole period.

1.4.2 Empirical Evidence from structural VAR

The estimated parameters of the standardized impact multiplier are shown in Table 1.2 (panel a).
Whilst residuals heteroscedasticity is a statistical tool to identify structural form shocks, ex post interpretation is obtained using the sign based restriction suggested in Table 1.1. Therefore, according to the sign based restriction, the first, second and third column show the standardized impact of a negative shock to credit demand, credit supply, and real economy, respectively. While the credit demand shock plays a bigger role than the credit supply shock on the loan interest rate, the reverse is true as for the impact on the loan growth rate. Although, on impact, the only statistically significant effect of an innovation to credit demand and credit supply is the one on the interest rate on loans (at 1 percent and 10 percent level of significance, respectively), results from Table 1.3 show a statistically significant cumulative effect of credit demand and credit supply shocks to both credit aggregates. Moreover, the empirical findings show that credit supply shocks play a more important role than those to credit demand in reducing the employment to population ratio. In particular, a one standard deviation shock to credit supply implies, on impact, a 1.3 percent change in the employment to population ratio (see Table 1.2, panel a) and a cumulative impact over a two-year horizon equal to 2.4 percent (see Table 1.3).
The real shock, interpreted as a negative one, due to its marginal depressing effect on the employment to population ratio, raises both the interest rate on loan and the growth in lending. The impact of the real shock on the employment rate is statistically significant over a two-year horizon.
Table 1.2 (panel b) shows that the identification assumption is satisfied because all the estimated parameters, λ_i, measuring the estimated relative variances, are distinct and

Figure 1.1: Boxplots for Loans interest rates, percent, for the Italian macro-areas, 2008-2014.

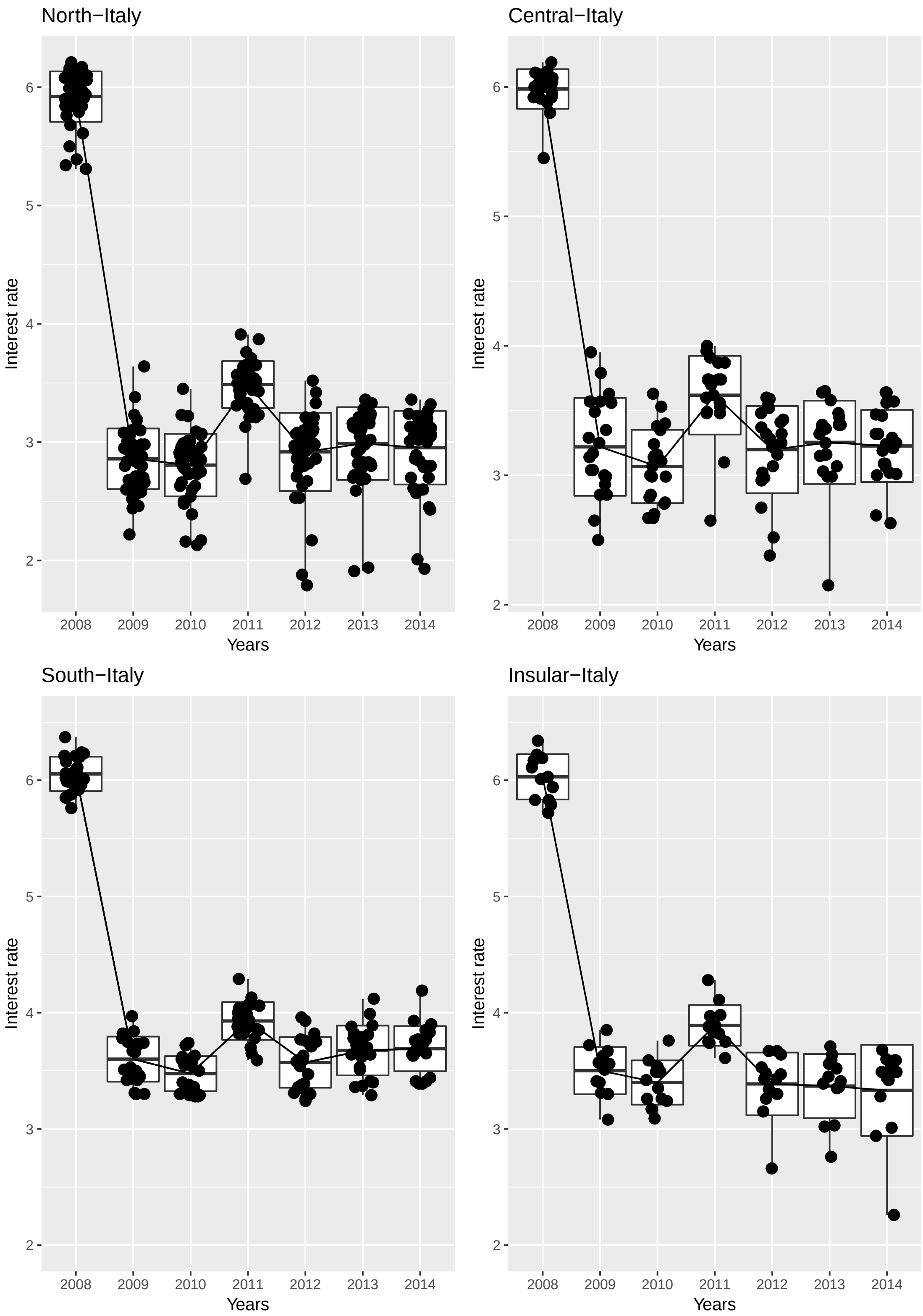


Figure 1.2: Boxplots for Loans growth rates, for the Italian macro-areas, 2008-2014.

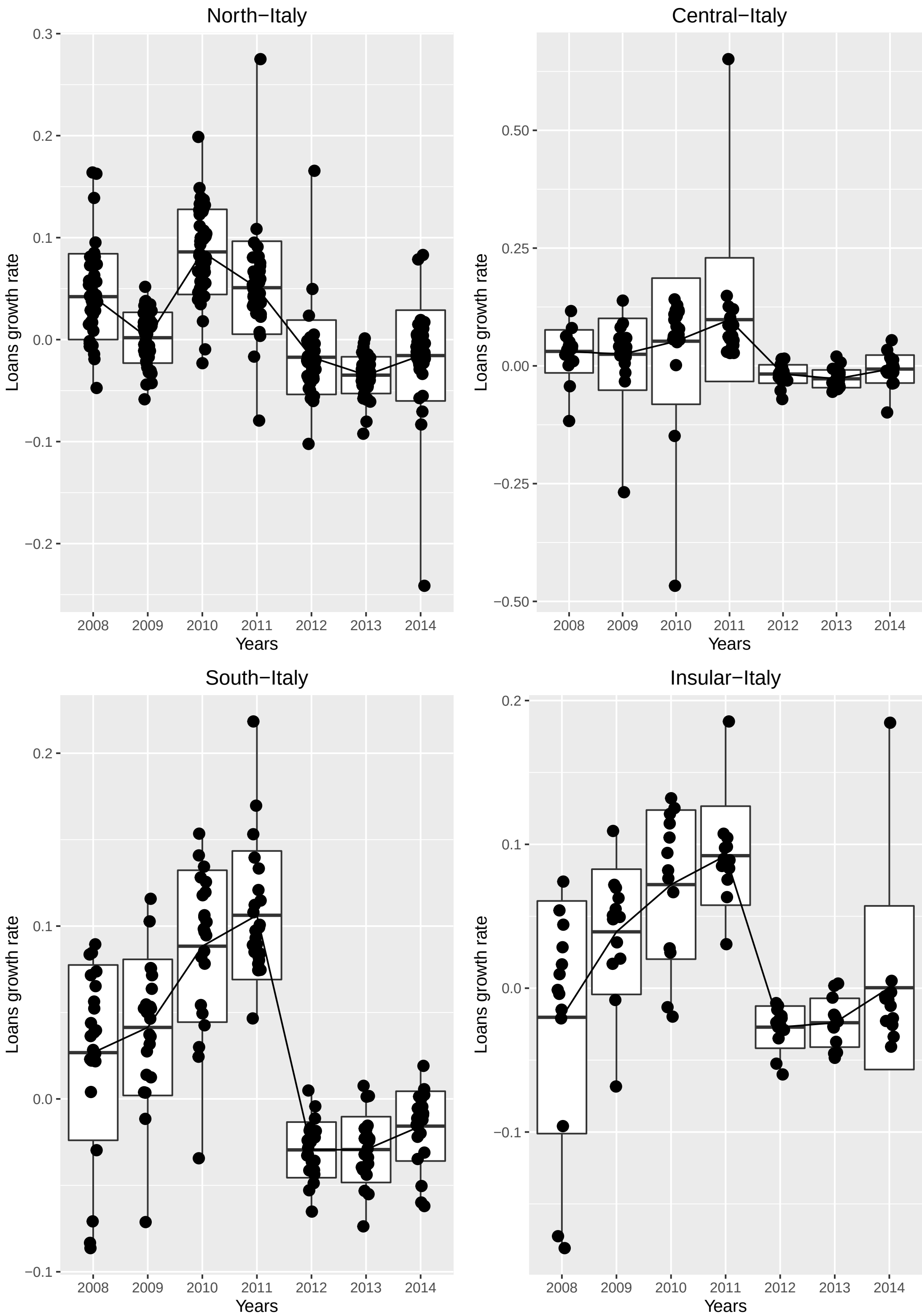
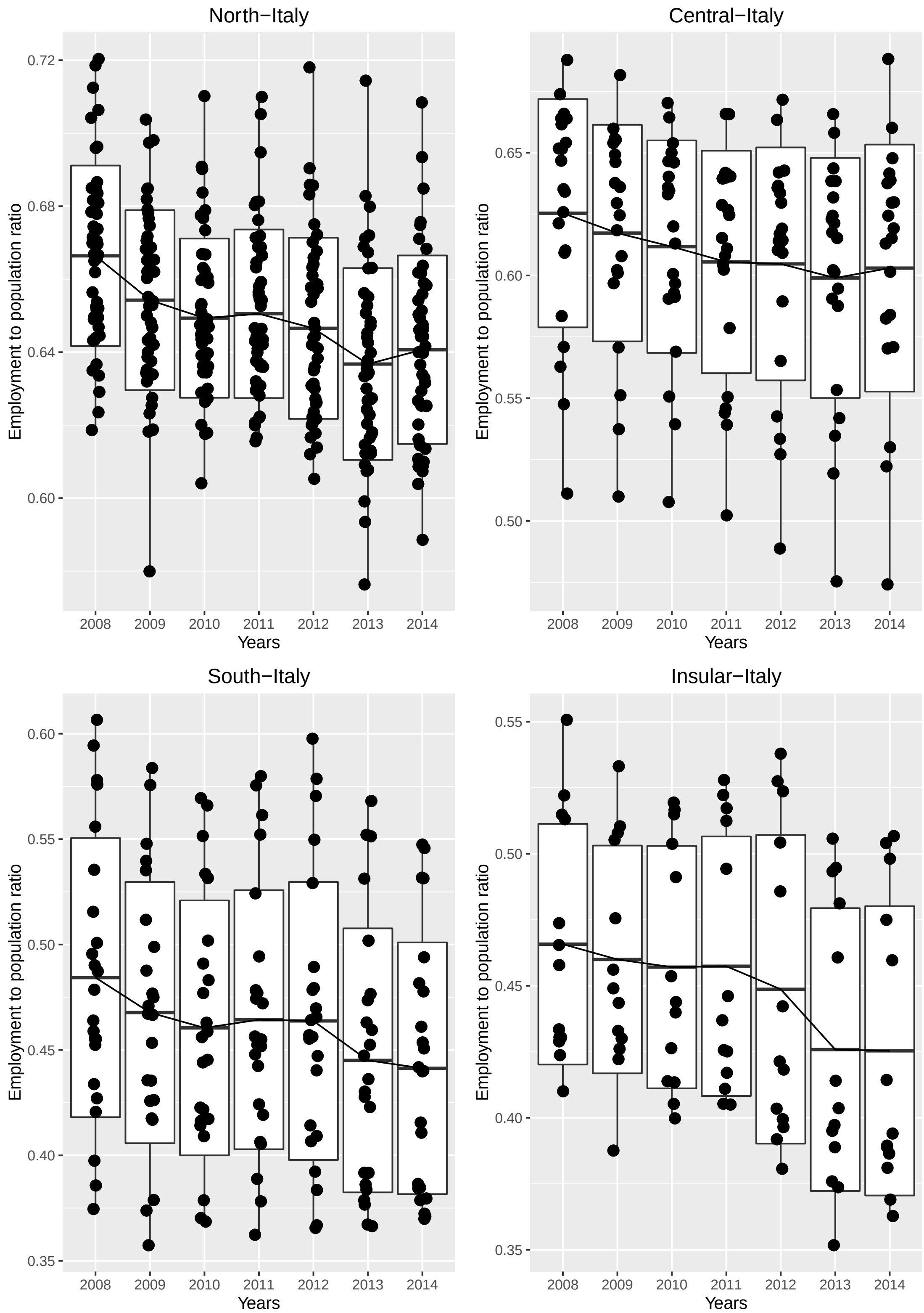


Figure 1.3: Boxplots for Employment to population ratio, for the Italian macro-areas, 2008-2014.



statistically significant.

The interpretation of the results in Table 1.2 (panel b) is based on the square root of the relative variances, in order to focus on the magnitude of shocks relative to the one for the North of Italy. The innovations hitting credit supply in Central, South and Insular Italy are all above unity suggesting that credit crunch hits the North of Italy less than the remaining macro-regions. In particular, the largest relative magnitude is observed for the credit supply shock hitting the South of Italy, and the magnitude of the innovation to credit supply in Central and Insular Italy is almost the same. While the largest relative magnitude of the credit demand shock observed is for Insular Italy, the South and Central Italy exhibit credit demand innovation with magnitude lower than the North. Finally, the largest magnitude of the real shock is observed in Central Italy (almost twice than the one for the North). Both South and Insular Italy exhibit a magnitude of the real shock above the corresponding one for the North (although much lower than the one for Central Italy).

Finally, Table 1.2 (panel c) shows the results of the over-identifying test restrictions using the LR statistic. The value of the log likelihood of the restricted model is equal to 2382.14, whilst the unrestricted log likelihood is equal to 2383.10. Therefore, the over-identifying restrictions are not rejected at 90 percent confidence level.

Following Lütkepohl (2011), I carry out with a historical decomposition (see Figures 1.4 and 1.5) in order to analyse the effects of credit market shocks on the real variable in the Italian provinces. My main focus is on the contribution of credit demand and supply shocks to the dynamic of the employment to population ratio (de-meaned at provincial level) for each macro-region. A credit demand shock seems to play a non-relevant role (with the exception of Insular-Italy) in explaining the downturn in the employment to population ratio.

I can observe, from historical decomposition, that credit supply shock plays an important role in tracking the dynamics of the employment rate in each macro-region, especially the slackening in employment rate in South and Insular Italy over 2013-2014.

To summarize, contrary to the empirical findings of Kick (2016), I find that credit supply shocks play a more important role than innovations to demand for credit for the dynamics of real economic activity in Italy. My results are in line with previous papers which focus on the credit crunch effect on real economy across Italian provinces (see Presbitero *et al.*, 2014; Barone *et al.*, 2016; Cingano *et al.*, 2016; Berton *et al.*, 2017). In particular, my findings about regional differences of credit crunch are in line with the ones of the study of Presbitero *et al.* (2014) which finds that the real economy of North Italy is more resilient to credit rationing, since, especially in the Southern regions, banks retracted disproportionally from markets that are more distant from their headquarters.

Table 1.2: Maximum Likelihood Estimation results of B and λ matrices.

Panel a: Standardized Impact multiplier (B matrix).

	Credit demand shock	Credit supply shock	Real shock
interest rate	-0.323^{****}	0.073^{*}	0.044^{*}
	(0.019)	(0.042)	(0.024)
$\Delta loans$	-0.001	-0.008	0.054^{****}
	(0.005)	(0.006)	(0.002)
empl. ratio	-0.004^{*}	-0.013^{****}	-0.002
	(0.002)	(0.001)	(0.001)

Panel b: Relative variances and magnitude of the shocks.

	Parameter	Magnitude
Central Italy		
Credit demand shock	0.933^{****}	0.966
Credit supply shock	1.238^{****}	1.113
Real shock	2.963^{****}	1.721
South Italy		
Credit demand shock	0.657^{****}	0.810
Credit supply shock	1.404^{****}	1.185
Real shock	1.234^{****}	1.111
Insular Italy		
Credit demand shock	1.405^{****}	1.185
Credit supply shock	1.207^{****}	1.099
Real shock	1.075^{****}	1.037

Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1

Panel c: Likelihood Ratio Test.

LR Test		Log-likelihood		Value
	Restricted model	2382.138	LR statistic	1.931
	Unrestricted model	2383.104	p-value	0.926

Note. All the parameters are estimated by ML. Asymptotic standard errors are provided in brackets. The relative variances (see panel b) are obtained setting to unity the elements on the first regime structural covariance matrix main diagonal, here refereed to the Northern Italy. The magnitude are obtained by taking the square root of the relative variances (e.g. the parameters in the second column).

Table 1.3: Cumulative Impact over a two year horizon.

	Mean	Lower bound	Upper bound
Credit demand shock			
interest rate	-0.291	-0.346	-0.262
$\Delta loans$	-0.008	-0.013	-0.002
empl. ratio	-0.010	-0.014	-0.004
Credit supply shock			
interest rate	0.123	0.066	0.168
$\Delta loans$	-0.013	-0.021	-0.005
empl. ratio	-0.024	-0.027	-0.022
Real shock			
interest rate	0.097	0.067	0.117
$\Delta loans$	0.060	0.055	0.066
empl. ratio	-0.004	-0.008	-0.001

Note. The First column is the mean value of bootstrapped distribution of the Cumulative Impact over a two year horizon, the last two columns are 16 percent and 84 percent bootstrapped confidence interval bounds.

Since my study shows that the Centre and South of Italy exhibit a relative higher magnitude of the credit supply shock, this contrasts the findings of Cingano *et al.* (2016) related to the territorial impact of rationing in lending. The authors find that the credit cut has been relatively homogeneous across borrowers and the firms with easier access to external finance or with a stronger liquidity position were more able to contain the negative consequences for investment (and, to less extent, on employment) of the drop in credit. Moreover, my findings contrast those from Barone *et al.* (2016) who find that the most severe credit rationing impact on real value added growth, during the recent financial crisis, occurred in the North and Central Italy which have firms relatively more dependent on external finance.

1.5 Conclusions

In this chapter, I have investigated the role of credit market shocks in explaining the downturn of the Italian economic activity using data at provincial level over 2008-2014. A number of studies of the Italian credit crunch are based on a two-stage estimation approach where in the first stage a credit supply indicator is identified through the Khwaja

Figure 1.4: Contribution of credit demand and supply shocks on historical decomposition of Employment rate, North and Central Italy, 2009-2014.

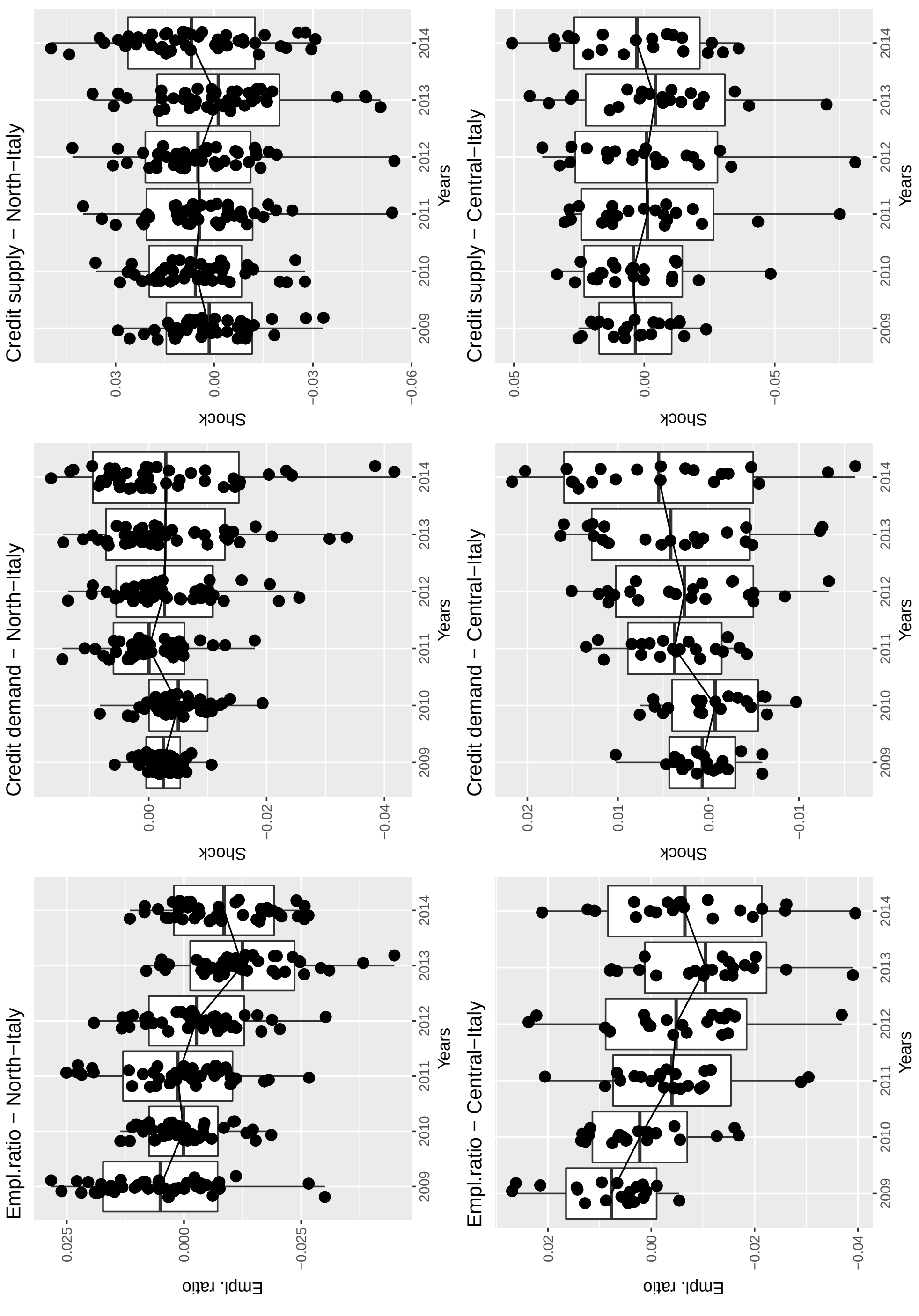
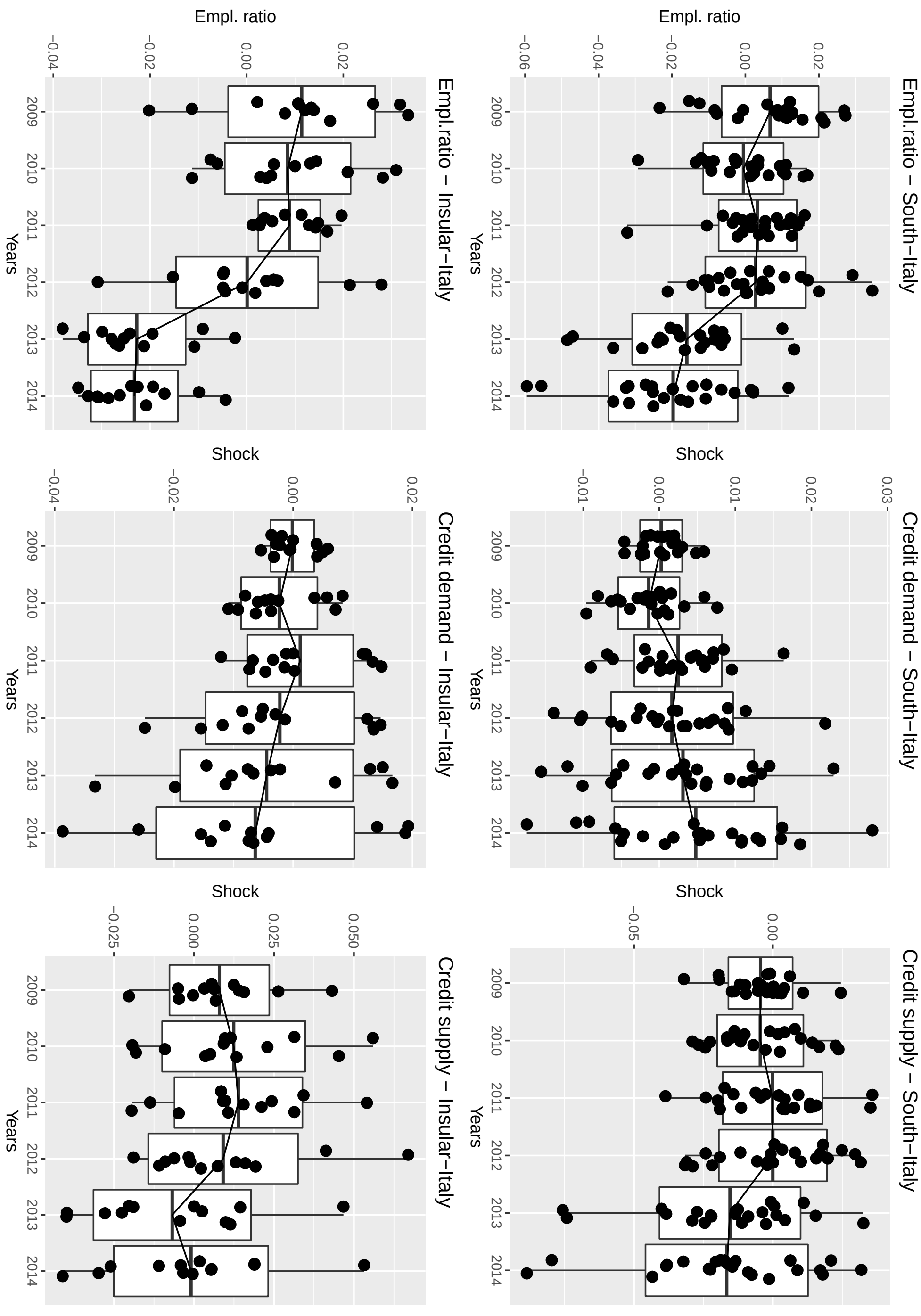


Figure 1.5: Contribution of credit demand and supply shocks on historical decomposition of Employment rate, South and Insular Italy, 2009-2014.



& Mian (2008) method which requires data on either bank-firms or bank-provinces relationships, observed in a pre and post crisis period. However, since my dataset is constrained only to a period of prolonged recession, my identification scheme is based on the changing variance of the structural shocks to a VAR fitted to interest rates, loans growth rates and employment ratio observed in the Italian macro-regions. Heteroscedasticity is only a statistical tool for the purpose of identification, therefore I have used ex post sign restrictions suggested by theory to identify demand and supply of credit shocks.

Differently from the empirical findings of Kick (2016), I find that credit supply shocks play a more important role than innovations to demand for credit. My findings related to a sizable and significant effect of credit supply on employment are in line with the studies, based on loans to Italian firms, of Barone *et al.* (2016), Cingano *et al.* (2016) and Berton *et al.* (2017).

Moreover, the empirical evidence shows that credit crunch hits the North of Italy less than the remaining macro-regions, especially the South-Italy. This findings are consistent with those of Presbitero *et al.* (2014) who find that the real economy of North Italy is more resilient to credit rationing, since, especially in the Southern regions, banks retracted disproportionally from markets that are more distant from their headquarters. An implication of these findings for Italy is that a key policy priority should therefore take into account the significant role of the credit supply. Taken together, these findings support the implementation of the recent Quantitative Easing adopted by the ECB to stimulate the economy.

Bibliography

Albertazzi, Ugo, & Marchetti, Domenico Junior. 2010. *Credit supply, flight to quality and evergreening: an analysis of bank-firm relationships after Lehman.* Temi di discussione (Working papers) No. 756. Bank of Italy.

Amiti, Mary, & Weinstein, David E. 2017. *How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data.* Journal of Political Economics, *forthcoming.*

Atta-Mensah, Joseph, & Dib, Ali. 2008. Bank lending, credit shocks, and the transmission of Canadian monetary policy. *International Review of Economics & Finance*, **17**(1), 159–176.

Barnett, Alina, & Thomas, Ryland. 2013. *Has weak lending and activity in the United Kingdom been driven by credit supply shocks?* Bank of England Working Paper No. 482. Bank of England.

Barone, Guglielmo, de Blasio, Guido, & Mocetti, Sauro. 2016. *The real effects of credit crunch in the Great Recession: evidence from Italian provinces.* Temi di Discussione (Working papers) No. 1057. Bank of Italy.

Berton, Fabio, Mocetti, Sauro, Presbitero, Andrea F, & Richiardi, Matteo. 2017. *Banks, firms and jobs.* Working paper No. 136. Mo.Fi.R.

Bijsterbosch, Martin, & Falagiarda, Matteo. 2014. *Credit supply dynamics and economic activity in Euro area countries: a time-varying parameter VAR analysis.* Working Paper Series No. 1714. European Central Bank.

Bofondi, Marcello, Carpinelli, Luisa, & Sette, Enrico. 2013. *Credit supply during a sovereign debt crisis.* Temi di discussione (Working papers) No. 909. Bank of Italy.

Bonaccorsi di Patti, Emilia, & Sette, Enrico. 2016. Did the securitization market freeze affect bank lending during the financial crisis? Evidence from a credit register. *Journal of Financial Intermediation*, **25**, 54–76.

Bottero, Margherita, Lenzu, Simone, & Mezzanotti, Filippo. 2015. *Sovereign debt exposure and the bank lending channel: impact on credit supply and the real economy.* Temi di discussione (Working papers) No. 1032. Bank of Italy.

Ciccarelli, Matteo, Maddaloni, Angela, & Peydro, Jose-Luis. 2015. Trusting the bankers: A new look at the credit channel of monetary policy. *Review of Economic Dynamics*, **18**(4), 979–1002.

Cingano, Federico, Manaresi, Francesco, & Sette, Enrico. 2016. Does credit crunch investment down? New evidence on the real effects of the Bank-Lending channel. *Review of Financial Studies*, **29**(10), 2737–2773.

Del Giovane, Paolo, Nobili, Andrea, & Signoretti, Federico M. 2017. Assessing the sources of credit supply tightening: Was the sovereign debt crisis different from Lehman? *International Journal of Central Banking*, **June 2017**.

Dörr, Sebastian, Raissi, Mehdi, & Weber, Anke. 2017. *Credit-supply shocks and firm productivity in Italy*. IMF Working paper No 17/67. International Monetary Fund.

ECB. 2011. *Recent developments in loans to the private sector*. Monthly Bulletin, January. European Central Bank.

Ehrmann, Michael, Fratzscher, Marcel, & Rigobon, Roberto. 2011. Stocks, bonds, money markets and exchange rates: measuring international financial transmission. *Journal of Applied Econometrics*, **26**(6), 948–974.

Greenstone, Michael, Mas, Alexandre, & Nguyen, Hoai-Luu. 2014. *Do credit market shocks affect the real economy? Quasi-experimental evidence from the Great Recession and "normal" economic times*. Working Paper No 20704. NBER.

Halvorsen, Jørn I, & Jacobsen, Dag Henning. 2014. How important can bank lending shocks be for economic fluctuations? *The North American Journal of Economics and Finance*, **29**, 104–123.

Hristov, Nikolay, Hülsewig, Oliver, & Wollmershäuser, Timo. 2012. Loan supply shocks during the financial crisis: Evidence for the Euro area. *Journal of International Money and Finance*, **31**(3), 569–592.

Khwaja, Asim Ijaz, & Mian, Atif. 2008. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *The American Economic Review*, **98**(4), 1413–1442.

Kick, Heinrich. 2016. *Spillover effects of credit demand and supply shocks in the EU countries: Evidence from a structural GVAR*. Unpublished manuscript.

Kilian, Lutz. 2013. Structural Vector Autoregressions. *Pages 515–54 of:* Hashimzade, N., & Thornton, M. (eds), *Handbook of Research Methods and Applications in Empirical Macroeconomics*.

Lanne, Markku, & Lütkepohl, Helmut. 2008. Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, **40**(6), 1131–1149.

Lütkepohl, Helmut. 2005. *New introduction to multiple time series analysis.* Springer-Verlag Berlin Heidelberg.

Lütkepohl, Helmut. 2011. *Vector autoregressive models.* Economics Working papers ECO 2011/30. European University Institute.

Lütkepohl, Helmut. 2012. *Identifying structural Vector Autoregressions via changes in volatility.* Discussion Papers 1259. DIW Berlin.

Lütkepohl, Helmut, & Netsunajev, Aleksei. 2015. *Structural Vector Autoregressions with heteroskedasticity - A comparison of different volatility models.* CESifo Working Paper Series 5308. CESifo Group Munich.

Mumtaz, Haroon, Pinter, Gabor, & Theodoridis, Konstantinos. 2015. *What do VARs tell us about the impact of a credit supply shock?* Working Papers 739. Queen Mary University of London, School of Economics and Finance.

Peersman, Gert. 2011. *Bank lending shocks and the Euro area business cycle.* Working papers 11/766. Ghent University, Faculty of Economics and Business Administration.

Presbitero, Andrea F, Udell, Gregory F, & Zazzaro, Alberto. 2014. The home bias and the credit crunch: A regional perspective. *Journal of Money, Credit and Banking*, **46**(s1), 53–85.

Rigobon, Roberto. 2003. Identification through heteroskedasticity. *Review of Economics and Statistics*, **85**(4), 777–792.

Chapter 2

Default rates spillovers: an analysis based on Italian regional data

2.1 Introduction

The aim of this chapter is the analysis of spatial spillover effects among 20 Italian regional default rates on loan granted to different categories of the private sector: consumer households, non-financial firms and producer households.[1]
The motivation of the analysis is due to the process of bank consolidation in Italy taking place during the 1990s, leading to a 33% reduction in the number of banks, from 1025 to 684, over the $1992 - 2013$ period (Papi *et al.*, 2015). The consolidation process was characterized by takeovers of the main distressed banks located in the Mezzogiorno (such as Banco di Napoli, Banco di Sicilia and other major savings banks) by Northern banks (mainly Unicredit and Intesa San Paolo).[2] Nowadays, the bulk of commercial banks located in Mezzogiorno are members of banking groups headquartered in the Northern part of the country (Zazzaro, 2006; Giannola *et al.*, 2013). More specifically, the study of Giannola *et al.* (2013) shows that, in 2010, more than 42% of branches operating in the Mezzogiorno were owned by banks headquartered outside the area and another 38% were attributable to banks which, whilst maintaining their headquarters in the Mezzogiorno, were part of banking groups whose parent bank was in the Center-North. The study

[1]According to the definition provided by Bank of Italy, producer households are defined as individual firms, informal partnership and unregistered company, producers of marketable goods and financial services with up to five employees; activities auxiliary to financial intermediation without employees.

[2]Mezzogiorno includes six Southern regions, such as Abruzzo, Apulia, Basilicata, Calabria, Campania and Molise, and the Islands of Sardinia and Sicily.

of Papi *et al.* (2015) shows, through network analysis, that the overall connectedness of geographical credit markets in Italy has significantly increased over time, whether measured at the provincial or regional level. Moreover, the authors confirm a growing centrality of few Northern Italian banking centers relegating the Southern credit markets and regions to the periphery. These findings support those in the study of Presbitero *et al.* (2014) showing an increasing functional distance (measured by the distance between bank branches and the bank headquarter) over recent years, hence a more striking core-periphery financial and banking divide. In particular, the headquarters of the large Northern banks will be less familiar with the local economic and social environment in the Mezzogiorno. As suggested by Alessandrini *et al.* (2009), physical distance between bank headquarters and local managers makes it difficult to gather and consequently report soft information to those higher up in the management chain and, consequently, monitor local managers. As a consequence, the allocation of decision-making power to local managers in the branches located in Mezzogiorno tends to decrease with distance. Therefore, one might expect a negative relationship between the credit growth and the distance between the centre and the periphery of the bank, especially during a crisis period characterized by credit tightening.

For this purpose, I use the Diebold-Yilmaz methodology, DY, based on the Generalized Forecast Error Variance Decomposition (GFEVD) (see Diebold & Yilmaz, 2012, 2014). The latter is obtained by employing the Adaptive Elastic net shrinkage estimator on a large Vector Autoregression (VAR) model, due to the 20 (region-specific) endogenous variables considered for each private sector category.

My study can be related to the one of Tola (2010) which is an application of the Pykhtin (2004) model to the Italian banking system to provide a static measure of concentration risk by industry sector and geographic region. For this purpose, the author uses a stationary multifactor structural Portfolio credit risk model, generating an unexpected loss measure that is in line with the Internal ratings-based (IRB) capital requirements. The use of the DY methodology, based on VAR estimation, is more suitable to address the evidence of non-stationarity I find in the proxies of default rates examined. Moreover, through a dynamic spillover analysis, using the DY methodology, I can assess whether there is evidence of an increase in the index of total default connectedness over $2011 - 2015$ (crisis period) relative to its long run value estimated by accounting for the whole sample period under investigation ($1996 - 2015$).

Finally, the DY methodology enables to retrieve indices of directional connectedness and, in particular, to assess whether the Mezzogiorno regions are more dependent (relative

to the Northern regions) on shocks arising from the other regions.[3] To detect macro-regional patterns in the spillover analysis, I use the approach proposed by Greenwood-Nimmo *et al.* (2015), say the GNS approach.
In my analysis, I use quarterly data for default rates on loans facilities to three categories of the private sector, that is consumer households, non-financial firms and producer households. The data, collected from the publicly-available Statistical Database of Bank of Italy, contain information on loan default rates for the 20 Italian regions, over the period $1996Q1 - 2015Q4$.

The results show an increase in the Total spillover index (hence there is evidence of a rise in spatial dependence) during the last observations of the sample ($2011Q4 - 2015Q4$), identified as a particular distressful period for the Italian economy.[4] These empirical findings are particularly striking for producer households. Using the approach proposed by Greenwood-Nimmo *et al.* (2015) (GNS), I find that the South and, to less extent, Northwest contribute the most to the financial stress of the other Italian macro-regions. Contrary to the South macro-region, the Islands financial distress largely depends on the others, especially consumer households and non-financial firms. Looking at the directional spillovers, I do not find evidence of a dependence from the North for all the Mezzogiorno regions. The dependence from North is only confirmed for the Islands, while shocks arising from South tend to largely spill over to both the Northwest and the Northeast.
This chapter is organized as follows. Section 2.2 reviews the literature on the pros and cons of bank geographical expansion. Section 2.3 describes the DY and the GNS approach on studying connectedness as well as the estimation procedure of a LASSO-VAR model. Section 2.4 describes data. Section 2.5 describes the empirical findings. Section 2.6 concludes.

2.2 Literature review

2.2.1 Geographical diversification: the evidence within country

Since 1990s, the Italian banking system has been characterized by a consolidation process which has largely involved a geographic expansion of Northern banks in Southern regions, through merger and acquisition (M&A) operations.

[3]There has been a growing number of applications of the DY methodology to financial institutions stock market returns and volatilities (see Diebold & Yilmaz, 2014; Demirer *et al.*, 2017, among the others). More recently, Cipollini *et al.* (2015) focused on volatility risk premia.

[4]The choice of the $2011Q4 - 2015Q4$ for the analysis of dynamic connectedness is motivated by the use of rolling regression in line with Diebold & Yilmaz (2012, 2014). Rolling estimation requires the use of a sufficient number of observations which in this study corresponds to a window size of 63 quarters.

Possible explanations might arise from the potential benefits of geographical diversification. In fact, as suggested by the traditional Portfolio Theory, geographical diversification/expansion is positively associated with a reduction in the risk related to a bank portfolio as longs as the different assets display low correlation (Goetz *et al.*, 2016). In particular, the authors find that a geographic diversification of bank's assets across Metropolitan Statistical Areas (MSAs) in the US diminishes a Bank Holding Company (BHC) risk. Using a geographic dispersion measure of deposits at branches level over the $1986 - 1997$ period, the authors also discover that the reduction of BHC's risk is positively associated to a geographic expansion when a BHC diversifies into MSAs that are economically different from its home MSA. In addition, a greater geographic diversification ought mitigate the adverse effects yielded by local business cycles. The study of Meslier *et al.* (2016) confirms the findings of Goetz *et al.* (2016), since there is evidence that (especially) small-size banks benefit from expanding geographically in non-contiguous markets with non-synchronized economic conditions. Consequently, a BHC may decide to extend its subsidiaries and branches across different areas in an attempt to reduce the exposure to its idiosyncratic local market risks. The study of Becchetti *et al.* (2014), focusing on 32 countries over the period $1998-2010$, shows that, in adverse phases of the business cycle, the share of loans to total assets of cooperative banks is higher than the one associated to other category of banks, with a positive effect on the growth of value added in the manufacturing sector and in those most dependent on external finance.

As for the European case, the study of Bonaccorsi di Patti *et al.* (2005) shows that, for Italy, the risk associated to poor geographical portfolio diversification can be particularly high during financial and economic downturns. The study of Illueca *et al.* (2013) highlights the negative effects of the portfolio risk concentration of Spanish banks, characterized by an ownership structure less geared to the attainment of economic performance, a focus on local community funding and an exposure toward the housing sector, particularly hit by the recent crisis.

Another strand of literature has questioned the attractiveness of geographic diversification, since the incentives to loan monitoring might be reduced, due to the difficulty in obtaining "soft-information".

Using data on commercial banks in Texas for the 1998, Brickley *et al.* (2003) suggest that a bank which extends its offer by opening branches and subsidiaries in distant areas ought face difficult in planning incentive-compensation for managers in the new branch, or subsidiary, arising the cost of monitoring their activity. Berger *et al.* (2005) point out that large BHCs which lend money to distant borrowers via their branches/subsidiaries tend to create weak relationships with the customers. By using survey data on small business lending over the 1994 – 1995 two-year period, the authors' results show that small banks have comparative advantages in supplying credit based on the "soft information". Moreover, as reported by the authors, there is evidence of a strong relationship between small banks and firms, and this can decisively reduce the probability of a borrower to be rationed. However, the authors find that local banks might be induced in funding obligors without paying attention to creditworthiness just to catch market shares.

The relevant role of local banks is also supported by the research of Berger & Udell (2002). The authors assert the importance of the relationship lending as well as suggesting that small banks might reduce the agency problems, generated by the accumulation of "soft information" by the loan officer, particularly when exogenous disturbances to credit market conditions, such as consolidation processes or changes in regulatory capital requirements, appear (see also Berger & Udell, 2006).

Imai & Takarabe (2011) focus on Japan and they examine how the nationwide city banks transmit large house price shocks to major city centre, intra-nationally, across geographical borders, to local economies in Japan. Presbitero *et al.* (2014) focus on Italy and they assess the role played by functional distance in the transmission mechanism of credit supply shocks across macro-regional economies.

As for the Italian evidence, using data on the asset and loan portfolio compositions of individual Italian banks during the 1993 - 1999 period, Acharya *et al.* (2006) find that diversification/expansion reduces bank returns as well as producing riskier loans, especially for high-risk banks. The study of Presbitero *et al.* (2014) highlights the negative effects of distance between the branches (or subsidiaries) and the BHC's headquarters. The authors find a positive causal relationship between the so-called "functional" distance, that is the distance between loan officer and banks' headquarters, and the tight of credit in Italy during the recent financial crisis. For the period of recession post-Lehman, Demma (2015) finds that, in Italy, local banks can mitigate the negative impact of the crisis on the quality of loans. Therefore, the benefits from soft information more than offset the effects due to adverse selection.

2.2.2 Geographical diversification: the evidence between countries

A number of studies have investigated the benefits of geographical expansion of large banks in advanced countries for the financial stability of emerging markets. The studies of Kaminsky & Reinhart (2000) and Van Rijckeghem & Weder (2001) were the first to identify a "common lender effects" as a cause of cross-border financial contagion. While the source of shock in the aforementioned studies was typically an emerging market, more recently the literature has also considered advanced countries as the originator of the crisis. This literature has concentrated on a "home bias" effect in credit allocation, implying that global banks exacerbate the transmission of financial shocks across regions, by moving funds from their peripheral to central (headquartered) markets. In particular, the international transmission of shocks may occur simply because internationally active banks suffer capital shortages in their domestic market (due to a crisis to the country where the headquarters are located) and they choose not to alter their portfolio mix of loans to domestic and foreign borrowers by cutting credit lines to both type of borrowers.

Cetorelli & Goldberg (2011) use BIS data on cross border lending and they focus on the capital flows reversals from developed to Emerging Asia, Latin America and Emerging Europe, right after the $2007-2008$ crisis period. The authors find that international banks contributed to the spreading of the crisis to emerging market economies. The major contribution of international banks to spreading the crisis was through a loan contraction manifesting through three separate channels: a contraction in direct, cross-border lending by foreign banks; a contraction in local lending by foreign banks' affiliates in emerging markets; and a contraction in lending supply by domestic banks as well, as a result of the funding shock to their balance sheet induced by the decline in interbank, cross-border lending.

Further evidence of a "flight to home" particularly striking during a $2007-2008$ (originated in the US) crisis period is provided by the study of Giannetti & Laeven (2012) which focuses on the syndicated loan market, a highly internationalized financial market, in which large banks lend to a variety of borrowers in a broad set of countries.

The crisis originator in the study of Schnabl (2012) is a liquidity shock originating in one country, Russia. The author, using both bank-to-bank lending and loan-level data, examines the role played by international banks to spreading the crisis in Peru. The author finds that the transmission is strongest for domestically-owned banks that borrow internationally, intermediate for foreign-owned banks, and weakest for locally funded banks. As argued by the author, the results suggest that lending between international banks establishes a transmission channel for bank liquidity shocks and that foreign bank ownership mitigates, rather than amplifies, the transmission through this channel.

Popov & Udell (2012) analyze the role played by global banks headquartered in Western Europe in spreading the $2007-2008$ crisis to Central and Eastern Europe. The authors find evidence that lending of multinational bank subsidiaries to firms located in these emerging markets was conditioned by the worsening in the balance sheet conditions of foreign parent banks.

The study of De Haas & Van Horen (2011) concentrates on the 118 largest banks in the cross-border syndicated loan market. In particular, the authors dataset allows to compare post-crisis and pre-crisis lending by each bank to each country. The authors find a strong and robust negative effect of geographical distance on lending stability, both in lending to advanced and to emerging markets. The authors find that banks that are further away from their customers are less reliable funding sources during a crisis. A second finding is that international banks with a local presence on the ground may be more stable providers of credit, that is foreign bank subsidiaries provide for a relatively stable credit source themselves, but their presence may also stabilise the cross-border component of bank lending.

2.3 Empirical methodology

2.3.1 The DY approach

Following Diebold & Yilmaz (2012, 2014), let me consider a K-multivariate covariance stationary process, $y_t = (y_{1t}, \ldots, y_{Kt})'$, described by a reduced form Vector Autoregression (VAR) model of order p:

$$y_t = \delta + A_1 y_{t-1} + \ldots + A_p y_{t-p} + u_t \tag{2.1}$$

where A_i, for $i = 1, \ldots, p$, are the $K \times K$ parameter matrices associated to the lagged variables, y_{t-i}, δ is a $K \times 1$ vector of constant terms and $u_t = (u_{1t}, \ldots, u_{Kt})' \sim N(0, \Sigma_u)$ is a vector of independent and identically distributed white noise disturbances, with a non-diagonal covariance matrix, $E(u_t u_t') = \Sigma_u$, which is not assumed to be diagonal.
Fixing $\delta = 0$, a stationary multivariate process admits the following Vector Moving Average representation of infinite order, VMA (∞):

$$y_t = \sum_{i=0}^{\infty} \Psi_i u_{t-i} \tag{2.2}$$

where Ψ_i, for $i = 1, \ldots, p$, are the $K \times K$ matrices of the VMA(∞) coefficients obtained from the following recursive substitution: $\Psi_i = A_1\Psi_{i-1} + A_2\Psi_{i-2} + \ldots + A_p\Psi_{i-p}$, with $\Psi_0 = I_K$ and $\Psi_i = 0$ for $i < 0$.[5]
From the reduced form VMA (∞), one can retrieve the impulse response function, which measures the time profile of a shock at time t on the expected value of the variables in the system after h periods, say $t + h$.
The studies of Diebold & Yilmaz (2012, 2014) follow the suggestions of Koop *et al.* (1996) and Pesaran & Shin (1998), relying on the generalized impulse response function

[5] See Lütkepohl (2005) and Diebold & Yilmaz (2012), for example.

which is not sensitive to the ordering of the variables as other identification scheme, such as the one based on the Cholesky decomposition of residuals covariance matrix (short-run restrictions).
Given a non decreasing information set, Ω_{t-1}, describing the known history of the economy before time t, Koop *et al.* (1996) and Pesaran & Shin (1998) define the generalized impulse response function (GIRF) of a variable at time $t+h$ hit by a shock a time t as follows:

$$GIRF(h, \eta, \Omega_{t-1}) = E(y_{t+h}|u_t = \eta, \Omega_{t-1}) - E(y_{t+h}|\Omega_{t-1}) = \Psi_h \eta \tag{2.3}$$

where η is a $K \times 1$ vector of shock, $\eta = (\eta_1, \ldots, \eta_K)'$, hitting the economy at time t and Ψ_h is the VMA(∞) coefficients matrix associated at time h. Therefore, the generalized impulse response can be seen as the difference between the expected value of a variable after h periods, conditional on shocks hitting the system at time t and the history up to $t-1$, and its expected value conditional on the previous history (defined as baseline profile). As suggested by Koop *et al.* (1996) and Pesaran & Shin (1998), an alternative approach consists of shocking the single j-th element of the vector of residuals, u_{jt}, for $j = 1, \ldots, K$, and comparing the expected value of a variable at time $t+h$ conditional on the j-th shock and the history of the system with the baseline profile:

$$GIRF(h, \eta_j, \Omega_{t-1}) = E(y_{t+h}|u_{jt} = \eta_j, \Omega_{t-1}) - E(y_{t+h}|\Omega_{t-1}) \tag{2.4}$$

Assuming a multivariate normal distribution of the residuals:

$$E(u_t|u_{jt} = \eta_j) = (\sigma_{1j}, \sigma_{2j}, \ldots, \sigma_{Kj})' \sigma_{jj}^{-1} \eta_j = \Sigma_u e_j \; \sigma_{jj}^{-1} \; \eta_j \tag{2.5}$$

where Σ_u is the covariance matrix of residuals in reduced form, σ_{jj} denotes the j-th main diagonal element entering Σ_u and e_j is a $K \times 1$ selection vector which takes value of 1 for the j-th element and zero elsewhere. The K-dimensional vector of generalized impulse responses to a shock arising from the j-th equation at time t after h periods is defined by combining eqs.(2.3), (2.4) and (2.5):

$$GIRF_j = \left(\frac{\Psi_h \Sigma_u e_j}{\sqrt{\sigma_{jj}}} \right) \left(\frac{\eta_j}{\sqrt{\sigma_{jj}}} \right) \tag{2.6}$$

or alternatively, by setting $\eta_j = \sqrt{\sigma_{jj}}$, it is possible to obtain the corresponding scaled version of the generalized impulse response function:

$$GIRF_j = \sigma_{jj}^{-\frac{1}{2}} \Psi_h \Sigma_u e_j \tag{2.7}$$

Under the assumption of normality of the residuals and linearity of the VAR model, Pesaran & Shin (1998) define the associated Generalized Forecast Error Variance Decomposition (GFEVD) matrix, $\mathcal{D}^H$, whose generic entry, d_{ij}^H, can be defined as follows:

$$d_{ij}^H = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' \Psi_h \Sigma_u e_j)^2}{\sum_{h=0}^{H-1} (e_i' \Psi_h \Sigma_u \Psi_h' e_i)} \tag{2.8}$$

and it measures the portion of the H-step ahead error variances in forecasting y_i due to shocks occurring to y_j, for $i, j = 1, \cdots, K$, such that $i \neq j$, Σ_u is the covariance matrix of the non-orthogonalized VAR residuals, u_t, σ_{jj} is the standard deviation of the error terms for the j-th equation, Ψ_h is the VMA(∞) coefficients matrix at time h and e_i, e_j are selection vectors with i-th and j-th element equal to unity and zero otherwise.
Since the shocks are not orthogonalized, the row sum of the entries in the variance decomposition matrix is not necessary equal to unity, $\sum_{j=1}^{K} d_{ij}^H \neq 1$. Therefore, Diebold & Yilmaz (2012, 2014) suggest a normalization by row sum of each element of the GFEVD matrix:

$$\tilde{d}_{ij}^H = \frac{d_{ij}^H}{\sum_{j=1}^{K} d_{ij}^H} \tag{2.9}$$

such that $\sum_{j=1}^{K} \tilde{d}_{ij}^H = 1$ and $\sum_{i,j=1}^{K} \tilde{d}_{ij}^H = K$, by construction.
The Connectedness table for the forecast horizon H is the GFEVD matrix augmented by a column containing the row sums of the off-diagonal elements of the GFEVD matrix and a row, where the column sums of the matrix off-diagonal entries take place. Finally, the average of all the off-diagonal elements appears, for $i \neq j$ (see Table 2.1).
The connectedness measures, both pairwise and system-wide, proposed by Diebold & Yilmaz (2012, 2014), can be retrieved directly from the Connectedness table. Each entry provides a *pairwise directional connectedness* measure from j to i:

$$C_{i \leftarrow j}^H = \tilde{d}_{ij}^H \tag{2.10}$$

For $i = j$, the pairwise measure explains the "*own share*" of the forecast error variance in a certain variable (e.g. a region) for a given forecast horizon. Generally, the GFEVD matrix ($\mathcal{D}^H$) is not symmetric, hence $C_{i \leftarrow j}^H \neq C_{j \leftarrow i}^H$.

Table 2.1: Connectedness Table

	y_1	y_2	$\cdots$	y_K	From others
y_1	$\tilde{d}^H_{11}$	$\tilde{d}^H_{12}$	$\cdots$	$\tilde{d}^H_{1K}$	$\sum_{j=1}^K \tilde{d}^H_{1j}, j \neq 1$
y_2	$\tilde{d}^H_{21}$	$\tilde{d}^H_{22}$	$\cdots$	$\tilde{d}^H_{2K}$	$\sum_{j=1}^K \tilde{d}^H_{2j}, j \neq 2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_K	$\tilde{d}^H_{K1}$	$\tilde{d}^H_{K2}$	$\cdots$	$\tilde{d}^H_{KK}$	$\sum_{j=1}^K \tilde{d}^H_{Kj}, j \neq K$
To others	$\sum_{i=1}^K \tilde{d}^H_{i1}$ $i \neq 1$	$\sum_{i=1}^K \tilde{d}^H_{i2}$ $i \neq 2$	$\cdots$	$\sum_{i=1}^K \tilde{d}^H_{iN}$ $i \neq K$	$\frac{1}{K}\sum_{i,j=1}^K \tilde{d}^H_{ij}$ $i \neq j$

Focusing on row and column sums, Diebold & Yilmaz (2012, 2014) propose the *Total* and *Directional connectedness* measures.
The sum of the GFEVD off-diagonal elements along each row of the Connectedness table, labelled *FROM* index, measures the *Directional connectedness from others to i-th element* of the table:

$$C^H_{i \leftarrow \bullet} = \sum_{\substack{j=1 \\ j \neq i}}^{K} \tilde{d}^H_{ij} \tag{2.11}$$

The index in eq.(2.11) measures the vulnerability (or the exposure) of a certain series to shocks originating in the remaining series for a given forecast horizon. Consequently, this index of directional connectedness can be interpreted as a measure of the vulnerability of series (e.g. regions) to systemic risk. The sum of the off-diagonal entries in the GFEVD matrix along each column, labelled *TO* index, measures, for a given forecast horizon, the *Directional connectedness of the j-th element to others*:

$$C^H_{\bullet \leftarrow j} = \sum_{\substack{i=1 \\ i \neq j}}^{K} \tilde{d}^H_{ij} \tag{2.12}$$

The index in eq.(2.12) measures the contribution of a shock occurring to a series (e.g. region) to the remaining series (e.g. regions).
Finally, the ratio between the sum of the off-diagonal entries in the GFEVD matrix and the sum of its total elements, that is simply the average of the off-diagonal entries in the GFEVD matrix, provides the *Total connectedness* index as:

$$C^H = \frac{1}{K} \sum_{\substack{i,j=1 \\ i \neq j}}^{K} \tilde{d}^H_{ij} \tag{2.13}$$

which is a measure of the inter-connectedness degree among different series (e.g. regions) for a given forecast horizon.

2.3.2 The GNS connectedness measures

For the purpose of interpretation of the results, I follow the approach recently proposed by Greenwood-Nimmo *et al.* (2015) which is based on constructing a block aggregation matrix from the GFEVD matrix, according to a certain aggregation scheme, arbitrarily defined.

In particular, given the K-dimensional vector of endogenous variables, the first step of the Greenwood-Nimmo *et al.* (2015) methodology (GNS) consists of re-normalizing the GFEVD matrix, such that $\mathcal{C}_R^H = K^{-1}\mathcal{D}^H$. The use of the re-normalization allows to obtain the connectedness measures, entering in $\mathcal{C}_R^H$, expressed as a portion of the total H-step forecast error variance (FEV) of the whole system.

After ordering (or re-ordering) the K endogenous variables, $y_t = (y_{1t}, \ldots, y_{Kt})'$, consistently to a selected scheme, it is possible to aggregate the endogenous variables into N groups. Since the generalized FEV approach is not sensitive to the ordering of the variables, the re-ordering procedure is not constrained to a particular scheme.

Suppose that the K endogenous variables are aggregated into N groups, where each n-th group contains a specific number of endogenous variables, K_n, with $n = 1, \ldots, N$. Greenwood-Nimmo *et al.* (2015) suggest to rewrite the above described $K \times K$ generalized forecast error variance decomposition (GFEVD) matrix at H-step ahead, D^H, as follows:

$$\underset{(K\times K)}{\mathcal{C}_R^H} = K^{-1}\begin{bmatrix} C^H_{1\leftarrow 1} & \cdots & C^H_{1\leftarrow K_1} & C^H_{1\leftarrow K_1+1} & \cdots & C^H_{1\leftarrow K_1+K_2} & \cdots & C^H_{1\leftarrow K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C^H_{K_1\leftarrow 1} & \cdots & C^H_{K_1\leftarrow K_1} & C^H_{K_1\leftarrow K_1+1} & \cdots & C^H_{K_1\leftarrow K_1+K_2} & \cdots & C^H_{K_1\leftarrow K} \\ C^H_{K_1+1\leftarrow 1} & \cdots & C^H_{K_1+1\leftarrow K_1} & C^H_{K_1+1\leftarrow K_1+1} & \cdots & C^H_{K_1+1\leftarrow K_1+K_2} & \cdots & C^H_{K_1+1\leftarrow K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C^H_{K_1+K_2\leftarrow 1} & \cdots & C^H_{K_1+K_2\leftarrow K_1} & C^H_{K_1+K_2\leftarrow K_1+1} & \cdots & C^H_{K_1+K_2\leftarrow K_1+K_2} & \cdots & C^H_{K_1+K_2\leftarrow K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C^H_{K\leftarrow 1} & \cdots & C^H_{K\leftarrow K_1} & C^H_{K\leftarrow K_1+1} & \cdots & C^H_{K\leftarrow K_1+K_2} & \cdots & C^H_{K\leftarrow K} \end{bmatrix} \tag{2.14}$$

where the n-th block, labelled as $\mathcal{C}_{n\leftarrow m}$, for $n, m = 1, \ldots N$, can be defined as:

$$\underset{(n\times m)}{\mathcal{C}^H_{n\leftarrow m}} = K^{-1}\begin{bmatrix} C^H_{\tilde{K}_n+1\leftarrow \tilde{K}_m+1} & \cdots & C^H_{\tilde{K}_n+1\leftarrow \tilde{K}_m+K_m} \\ \vdots & \ddots & \cdots \\ C^H_{\tilde{K}_n+K_n\leftarrow \tilde{K}_m+1} & \cdots & C^H_{\tilde{K}_n+K_n\leftarrow \tilde{K}_m+K_m} \end{bmatrix} \tag{2.15}$$

where $\tilde{K}_n = \sum_{n=1}^{n-1} K_n$[6]. Therefore, the GFEVD matrix can be represented as a block matrix, one for each of the N groups:

$$\underset{(K\times K)}{\mathcal{C}_R^H} = \begin{bmatrix} \mathcal{C}_{1\leftarrow 1}^H & \mathcal{C}_{1\leftarrow 2}^H & \cdots & \mathcal{C}_{1\leftarrow N}^H \\ \mathcal{C}_{2\leftarrow 1}^H & \mathcal{C}_{2\leftarrow 2}^H & \cdots & \mathcal{C}_{2\leftarrow N}^H \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_{N\leftarrow 1}^H & \mathcal{C}_{N\leftarrow 2}^H & \cdots & \mathcal{C}_{N\leftarrow N}^H \end{bmatrix} \tag{2.16}$$

As stated by Greenwood-Nimmo *et al.* (2015), the blocks lying on the diagonal of $\mathcal{C}_R^H$ in eq.(2.16), that is the $\mathcal{C}_{n\leftarrow n}^H$ matrices, provide information on the within-group FEV contributions. For the n-th group, the *Total within-group* FEV contribution is computed as follows:

$$\mathcal{W}_{n\leftarrow n}^H = \mathbf{1}'_{K_n} \mathcal{C}_{n\leftarrow n}^H \mathbf{1}_{K_n} \tag{2.17}$$

where $\mathbf{1}_{K_n}$ is a $K_n \times 1$ vector of ones. The *Total within-group* measures the contribution of the variables entering a group to its own H-step ahead FEV (see also Park & Shin, 2017). The off-diagonal blocks entering in $\mathcal{C}_R^H$, that is the $\mathcal{C}_{n\leftarrow m}^H$ matrices, with $n \neq m$, provide information on the spillover effects among two different groups. In a similar fashion to the pairwise connectedness measures proposed in the DY approach, Greenwood-Nimmo *et al.* (2015) define the spillover effect *from* group m to group n as:

$$\mathcal{F}_{n\leftarrow m}^H = \mathbf{1}'_{K_n} \mathcal{C}_{n\leftarrow m}^H \mathbf{1}_{K_m} \tag{2.18}$$

while the spillover effect *to* group m from group n as:

$$\mathcal{T}_{m\leftarrow n}^H = \mathbf{1}'_{K_m} \mathcal{C}_{m\leftarrow n}^H \mathbf{1}_{K_n} \tag{2.19}$$

It is important to note that $\mathcal{F}_{n\leftarrow m}^H$ and $\mathcal{T}_{n\leftarrow m}^H$ coincide.
Furthermore, Greenwood-Nimmo *et al.* (2015) provide a set of "system-wide" connectedness measures. In particular, the total *From*, *To* and *Net* contributions for group n can be defined as follows:

[6] As discussed by Greenwood-Nimmo *et al.* (2015), the number of variables for each group can be different among groups.

$$\mathcal{F}^H_{n\leftarrow\bullet} = \sum_{m=1,m\neq n}^{N} \mathcal{F}^H_{n\leftarrow m} \quad , \quad \mathcal{T}^H_{\bullet\leftarrow n} = \sum_{m=1,m\neq n}^{N} \mathcal{T}^H_{m\leftarrow n} \quad \text{and} \quad \mathcal{N}^H_{\bullet\leftarrow n} = \mathcal{T}^H_{\bullet\leftarrow n} - \mathcal{F}^H_{n\leftarrow\bullet} \tag{2.20}$$

where $\mathcal{F}^H_{n\leftarrow\bullet}$ measures the contribution to the FEV of the n-th group from the rest of the system, $\mathcal{T}^H_{\bullet\leftarrow n}$ measures the contribution of the n-th group to the FEV of the remaining groups and $\mathcal{N}^H_{\bullet\leftarrow n}$ measures to what extent the n-th group is a net transmitter or receiver of spillover effects.[7]

Finally, Greenwood-Nimmo *et al.* (2015) introduce two additional measures of connectedness: the *Dependence* and the *Influence* index. The *Dependence* index ($\mathcal{O}_n$) measures to what extent the n-th group is affected by external conditions:

$$\mathcal{O}^H_n = \frac{\mathcal{F}^H_{n\leftarrow\bullet}}{\mathcal{W}^H_{n\leftarrow n} + \mathcal{F}^H_{n\leftarrow\bullet}} \tag{2.21}$$

with $0 \leq \mathcal{O}^H_n \leq 1$. In particular, the role of external shocks in the conditions of group n decreases as $\mathcal{O}^H_n$ tends to zero, while the importance of external conditions increases as $\mathcal{O}^H_n$ becomes closer to one. The *Influence* index ($\mathcal{I}_n$) provides a measure of the role played by group n as influencer of the system:

$$\mathcal{I}^H_n = \frac{\mathcal{N}^H_{\bullet\leftarrow n}}{\mathcal{T}^H_{\bullet\leftarrow n} + \mathcal{F}^H_{n\leftarrow\bullet}} \tag{2.22}$$

with $-1 \leq \mathcal{I}^H_n \leq 1$. The use of the *Influence* index allows to determine whether the n-th group is a net shock recipient ($-1 \leq \mathcal{I}^H_n < 0$), a net shock transmitter ($0 < \mathcal{I}^H_n \leq 1$) or neither of the two roles ($\mathcal{I}^H_n = 0$) (see Greenwood-Nimmo *et al.*, 2015).

2.3.3 Estimation procedure

2.3.3.1 Shrinkage estimators

Given a relatively large number of endogenous variables ($K = 20$) in the VAR model, I use a Lasso-VAR approach where the current values of the K endogenous variables are considered as dependent variables and their lagged values are treated as explanatory

[7]Furthermore, Greenwood-Nimmo *et al.* (2015) define other two aggregate connectedness measures which can be derived from eqs.(2.17) and (2.20). The former (labelled aggregate "Heatwave" index) $\mathcal{H}^H = \sum_{n=1}^N \mathcal{W}^H_{n\leftarrow n}$, provides a measure of the importance of own (local) conditions for the whole system, while the latter (aggregate "Spillover" index), $\mathcal{S}^H = \sum_{n=1}^N \mathcal{F}^H_{n\leftarrow\bullet} \equiv \sum_{n=1}^N \mathcal{T}^H_{\bullet\leftarrow n}$, captures the magnitude of spillover effects among groups. Note that $\mathcal{H}^H + \mathcal{S}^H = 1$ and $\sum_{n=1}^N \mathcal{N}^H_{\bullet\leftarrow n} = 0$, by construction.

variables (Hsu *et al.*, 2008; Davis *et al.*, 2016).

The LASSO (Least Absolute Shrinkage and Selection Operator) regularization technique was originally introduced by the research of Tibshirani (1996). The LASSO, which provides estimation and variable selection, is particularly attractive when the unknown parameters are greater than the number of observations. In such as context, the LASSO shrinks the coefficient to exact zero, generating sparsity in the model representation.

In linear regression models, considering a vector of responses, $y_t = (y_1, \ldots, y_T)' \in \mathbb{R}$, and K independent variables, $x_{jt} = (x_{j1}, \ldots, x_{jT})' \in \mathbb{R}^K$, with $j = 1, \ldots, K$, the LASSO estimator solves the following convex optimization problem:

$$\hat{\beta}_{LASSO} = \underset{(\beta_0, \beta_j) \in \mathbb{R}^{K+1}}{\arg\min} \left\{ \sum_{t=1}^{T} \left(y_t - \beta_0 - \sum_{j=1}^{K} \beta_j x_{jt} \right)^2 \right\} \quad \text{subject to} \quad \sum_{j=1}^{K} |\beta_j| \leq c \tag{2.23}$$

Alternatively, using the Lagrange multiplier, one can write eq.(2.23) as follows:

$$\hat{\beta}_{LASSO} = \underset{(\beta_0, \beta_j) \in \mathbb{R}^{K+1}}{\arg\min} \quad \|y - \beta_0 - \sum_{j=1}^{K} \beta_j x_j\|_{\ell_2}^2 + \lambda \|\beta_j\|_{\ell_1} \tag{2.24}$$

where $\|y - \beta_0 - \sum_{j=1}^{K} \beta_j x_j\|_{\ell_2}^2 = \|u\|_{\ell_2}^2 = (\sqrt{\sum_{t=1}^{T} u_t^2})^2$ is the square of the Euclidean norm of the vector u, while the second part of the minimization problem is the ℓ_1-norm, that is $\|\beta_j\|_{\ell 1} = \sum_{j=1}^{K} |\beta_j|$. Furthermore, $c \geq 0$, or alternatively $\lambda \geq 0$, is a tuning parameter which controls the amount of shrinkage (Tibshirani, 1996).

Although the LASSO estimation procedure has seen a large number of applications in literature during the last two decades, it has also been criticized by some authors.

For example, it has been argued that LASSO does not perform well in terms of prediction power when the variables are highly correlated (Tibshirani, 1996). Furthermore, as Zou & Hastie (2005) point out, in case of high correlation among variables, LASSO does not encourage group selection, that is if two or more variables display high correlation, a selection method should include the whole group whether one of those variables is selected.

To this end, Zou & Hastie (2005) propose the so-called Elastic net (ENET) estimator which solves the following optimization problem:

$$\hat{\beta}_{ENET} = \underset{(\beta_0, \beta_j) \in \mathbb{R}^{K+1}}{\arg\min} \quad \|y - \beta_0 - \sum_{j=1}^{K} \beta_j x_j\|_{\ell_2}^2 + \lambda \sum_{j=1}^{K} \left[\alpha |\beta_j| + (1 - \alpha) \beta_j^2 \right] \tag{2.25}$$

where the elastic net penalty, $\alpha\|\beta\|_{\ell_1} + (1-\alpha)\|\beta\|^2_{\ell_2}$, is a convex combination of the ℓ_1-norm (LASSO) and ℓ_2-norm (Ridge regression).[8] Whether $\alpha = 1$, the elastic net penalty becomes the LASSO penalty. Oppositely, fixing $\alpha = 0$, the penalty turns into the Ridge regression. In particular, according to Zou & Hastie (2005), the ℓ_1-norm ensures automatic variables selection and shrinkage, simultaneously, while the Ridge regression's penalty encourages group selection, improving the prediction power of the estimator.
Moreover, Fan & Li (2001) argue that the LASSO estimator does not simultaneously respect the so-called oracle-properties, that is an ideal penalized least square procedure must *i)* identify the correct model whenever the right regularization parameter is chosen (consistency in variable selection), and *ii)* it has an asymptotically normal distribution[9].
Zou (2006) proposes an alternative version of the LASSO estimator, the Adaptive LASSO (ALASSO), where different weights are used for the penalization of each coefficient. The ALASSO is the estimator which solves the following convex optimization problem with the ℓ_1 penalty:

$$\hat{\beta}_{ALASSO} = \underset{(\beta_0,\beta_j)\in\mathbb{R}^{K+1}}{\arg\min} \quad \|y - \beta_0 - \sum_{j=1}^{K}\beta_j x_j\|^2_{\ell_2} + \lambda\sum_{j=1}^{K}\hat{w}_j|\beta_j| \tag{2.26}$$

where $\hat{w}_j$ is a vector of j "adaptive" weights. In literature, the weights are generally defined as $\hat{w}_j = 1/|\hat{\beta}_j|^\gamma$, where $\hat{\beta}$ is a root-n-consistent estimator of β and $\gamma > 0$. As reported in Zou (2006), under specific conditions, that is the weights are data-dependent and suitably defined, the ALASSO estimator is consistent in choosing the right subset of variables and asymptotically normal. Therefore, differently from the Elastic net, the Adaptive LASSO estimator respects the oracle properties.
Nevertheless, the ALASSO penalization does not achieve the performance in terms of stability of the Elastic net. For this reason, Zou & Zhang (2009) propose an alternative penalization which combines the Adaptive LASSO penalization and the ridge regression, the Adaptive Elastic net (AdaEnet). The resulting estimator is defined as follows:

$$\hat{\beta}_{AdaEnet} = \underset{(\beta_0,\beta_j)\in\mathbb{R}^{K+1}}{\arg\min} \quad \|y - \beta_0 - \sum_{j=1}^{K}\beta_j x_j\|^2_{\ell_2} + \lambda\sum_{j=1}^{K}\Big[\alpha\ \hat{w}_j|\beta_j| + (1-\alpha)\beta_j^2\Big] \tag{2.27}$$

where the adaptive weights are generally constructed as $\hat{w}_j = 1/\hat{\beta}^\gamma_{Enet}$, with $\gamma > 0$. As demonstrated by Zou & Zhang (2009), the Adaptive Elastic net has the oracle properties and, at the same time, the use of the ℓ_2 penalty provides stability in case of

[8] The expression in eq.(2.25) refers to what Zou & Hastie (2005) define the naïve elastic net, which is then rescaled to obtain the elastic net estimator (see also Zou & Zhang, 2009).
[9] Cfr. also Zou (2006) for a further explanation of the oracle properties.

high-dimensional data.

2.3.3.2 LASSO-VAR(1) model

Since the K-dimensional time series are not stationary, I estimate a sparse VAR(1) model fitted to the first order difference of the logit transformation of the loan default rates, by using the Adaptive Elastic net estimator proposed by Zou & Zhang (2009). In particular, given a K-dimensional vector of time series, $y_t = (y_{1t}, \ldots, y_{Kt})'$, the model has the following reduced form representation:

$$\Delta y_t = \delta + A_1 \Delta y_{t-1} + u_t \tag{2.28}$$

where A_1 is the $K \times K$ coefficients matrix of the lagged variables, Δy_{t-1}, δ is a $K \times 1$ vector of constant terms and $u_t \sim N(0, \Sigma_u)$ are the white-noise disturbances with a non-singular covariance matrix, $E(u_t u_t') = \Sigma_u$, which is not assumed to be diagonal.
Recently, a large number of researchers have shown the attractiveness of estimating the sparse VAR process through the estimation of K separate equations (see Kock & Callot, 2015; Demirer *et al.*, 2017). In line with this strand of literature, I carry out with an equation-by-equation VAR estimation by using the version of the Adaptive Elastic net used in the study of Demirer *et al.* (2017), which solves the following optimization problem for each of the K equations:

$$\hat{\beta}_{k,AdaEnet} = \underset{(\delta,\beta_j)\in\mathbb{R}^{K+1}}{\arg\min} \; \|\Delta y_t - \delta - \sum_{j=1}^{K} \beta_j \Delta y_{jt-1}\|_{\ell_2}^2 + \lambda \sum_{j=1}^{K} \hat{w}_j \left[\alpha|\beta_j| + (1-\alpha)\beta_j^2\right] \tag{2.29}$$

where β_j, $j = 1, \ldots, K$, is the j-th row vector of the $K \times K$ coefficient matrix, A_1, and $\hat{w}_j = 1/|\hat{\beta}_{j,OLS}|^\gamma$, with $\gamma = 1$, is the j-dimensional vector of weights. In order to estimate the model, I fix $\alpha = 0.5$ and I select the tuning parameter, λ, by applying a 10-fold cross validation equation by equation, separately (see also Bonaldi *et al.*, 2015).[10] As stated by Demirer *et al.* (2017), the use of a LASSO-based estimator produces sparsity on the coefficient matrix, however no sparsity is imposed on the resulting covariance matrix of VAR residuals.[11]

[10] The computational analysis is run by using the **glmnet** package in R developed by Friedman *et al.* (2010), which uses algorithms based on cyclical coordinate descent methods. I allow the **glmnet** package to standardize the covariates, that is $\frac{1}{T}\sum_{t=1}^{T} x_{jt} = 0$ and $\frac{1}{T}\sum_{t=1}^{T} x_{jt}^2 = 1$. Once standardizing the variables, the **glmnet** package always returns the coefficients to the original scale, automatically.

[11] See Demirer *et al.* (2017) for further details.

Once the sparse VAR(1) model is estimated, I construct the GFEVD matrix, $\mathcal{D}^H$, with a the generic entry defined as follows:

$$d_{ij}^H = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' \Psi_h^* \Sigma_u e_j)^2}{\sum_{h=0}^{H-1} (e_i' \Psi_h^* \Sigma_u \Psi_h^{*\prime} e_i)} \tag{2.30}$$

Since the endogenous variables enter the model in their first order difference, I construct the GFEVD by computing the cumulative Moving Average (MA) coefficients matrices, Ψ_h^*, at forecast horizon h. In my analysis, all the connectedness measures retrieved from the GFEVD are computed by considering a four-quarter forecast horizon ($H = 4$).

2.4 Data

I use data for default rates on loans to three categories of the private sector, that is consumer households (consumers), non-financial firms (nfi) and producer households (producers), in the 20 Italian regions, for a total of 4800 observations. According to the definition provided by Bank of Italy, the default rate on loans in a certain quarter t is the ratio between the amount of credit used by borrowers who become "adjusted bad debtors" during the observed quarter t and the amount of credit used by all the borrowers, not classified as "adjusted bad debtors" by the Central Credit Register, at the end of the previous quarter, $t-1$ (see also Bofondi & Gobbi, 2004).

The dataset, collected from the publicly-available database of Bank of Italy, includes quarterly frequency observations over the period $1996Q_1 - 2015Q_4$.[12] In my analysis, I use the NUTS1 and NUTS2 classifications imposed by the European Commission. For the Italian case, the former comprises of 5 groups of regions (or macro-regions), while the latter refers to the 20 regions (see Table 2.2).

Figures 2.1, 2.2 and 2.3 show the $K = 20$ regional loan default rate time series for each of the three categories of the private sector.

In general, the loan default rates reported by the Southern and Insular regions exhibit the highest values over the whole observed period. The loan default rates for consumer households (see Figure 2.1) show a decreasing trend with low values of the ratio reported in the last part of the sample. The loan default rates for non-financial firms show a rising pattern, especially the ones reported by the Northern and Central regions, over the last quarters (see Figure 2.2). Finally, Figure 2.3 shows that the loans default rate series for producer households tend to remain steady in the most of the Northern and Central regions, with the exception of Lazio, while there is evidence of a decline in the value of

[12] The missing values in the default rate on loan facilities series are replaced by linear interpolation.

Figure 2.1: Default rates on loan facilities (in percentage) for Consumer households in the Italian regions, from 1996$Q1$ to 2015$Q4$.

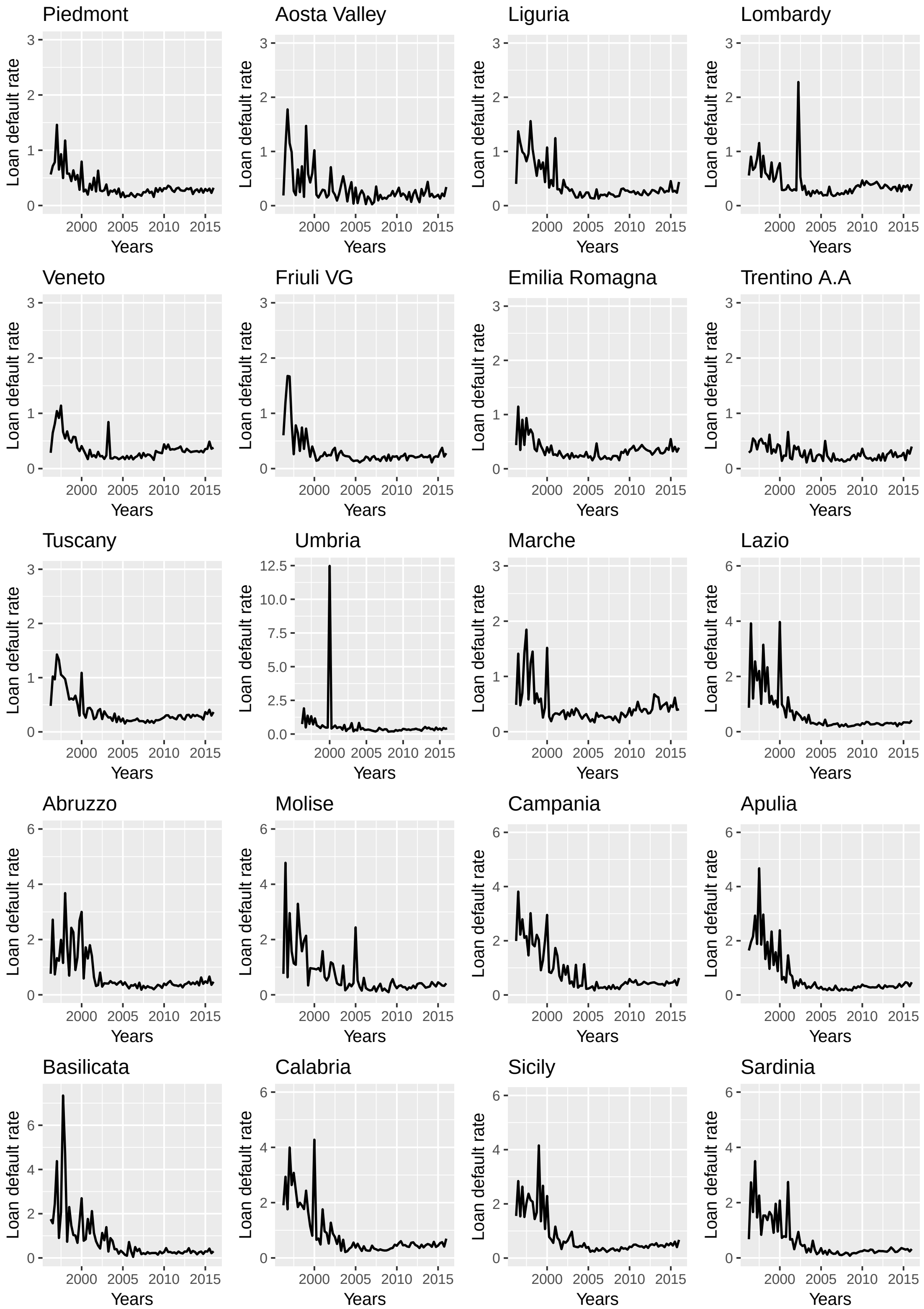


Figure 2.2: Default rates on loan facilities (in percentage) for Non-financial firms in the Italian regions, from 1996$Q1$ to 2015$Q4$.

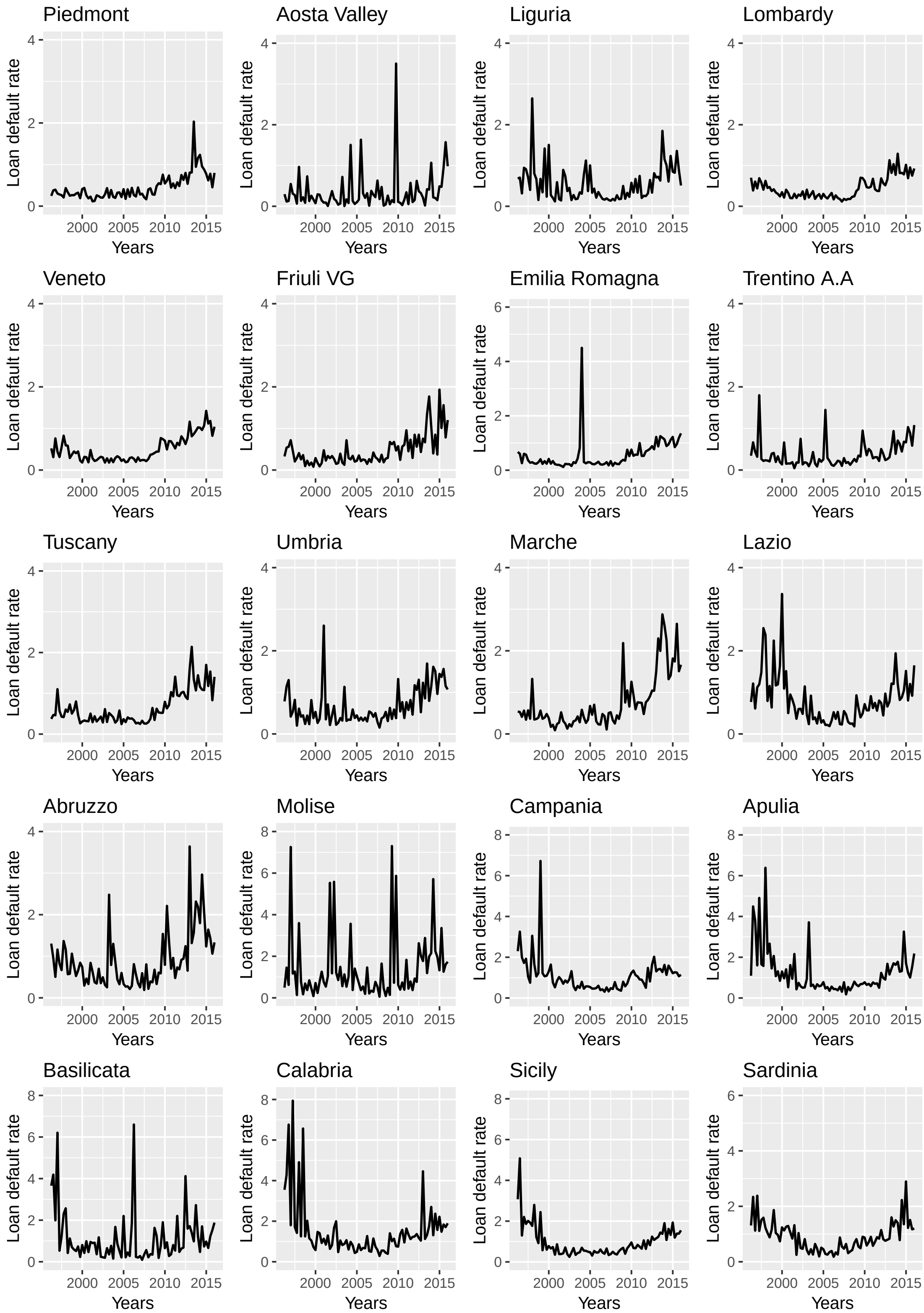


Figure 2.3: Default rates on loan facilities (in percentage) for Producer households in the Italian regions, from 1996$Q1$ to 2015$Q4$.

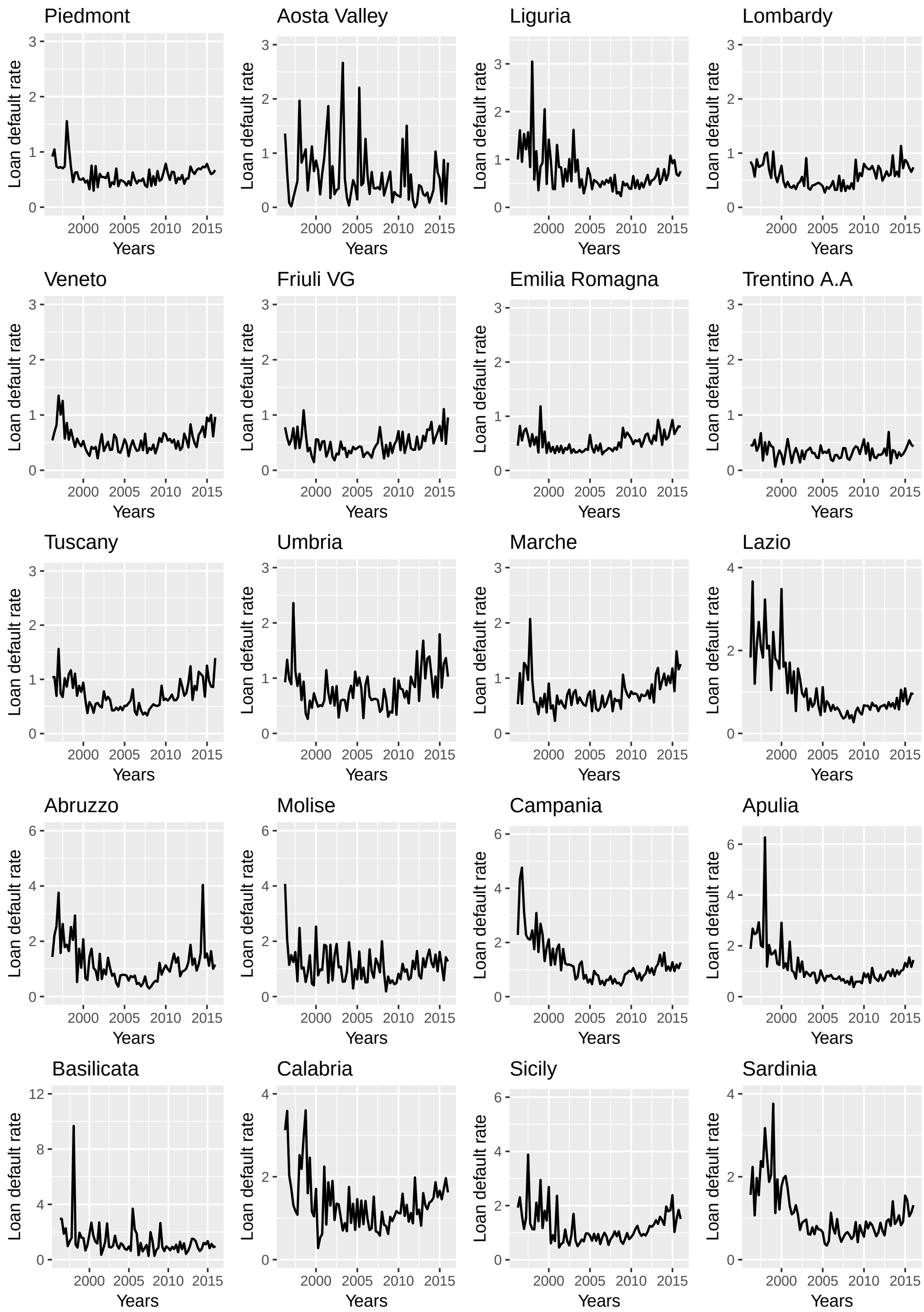


Table 2.2: Italian regions grouped at NUTS 1 (macro-regional) level.

Northwest	Northeast
Aosta Valley	Emilia-Romagna
Liguria	Friuli Venezia-Giulia
Lombardy	Trentino Alto-Adige
Piedmont	Veneto
Centre	**South**
Lazio	Abruzzo
Marche	Apulia
Tuscany	Basilicata
Umbria	Calabria
	Campania
Islands	Molise
Sardinia	
Sicily	

Note. Mezzogiorno includes the Southern regions and the islands of Sardinia and Sicily.

the ratio reported by some Southern regions.

Following Virolainen (2004), Foglia *et al.* (2009) and Guarda *et al.* (2012), I apply the logit transformation to the loan default rate series:

$$y_{ikt} = ln\left(\frac{p_{ikt}}{1 - p_{ikt}}\right) \tag{2.31}$$

where p_{ikt} is the default rate on loan facilities reported in the i-th category of the private sector, for the k-th variable (region) at time t. Since the loan default rate, p_{ikt}, ranges in the interval $[0, 1]$, the "logit transformation" in eq.(2.31) extends the boundary, moving to an unconstrained space of values, $y_{ikt} \in [-\infty, +\infty]$.

Table 2.3 shows the results of the Augmented Dickey-Fuller (ADF) test for the presence of unit roots in the time series under investigation. According to the Dickey-Fuller critical values, the null hypothesis, that is the time series are not stationary, is not rejected for almost all the time series. The non-stationarity is also confirmed by the use of the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root test proposed in Kwiatkowski *et al.* (1992), which tests for the null hypothesis that the series is level or trend stationary (see Table 2.4). Therefore, I fit a VAR model to the first order difference of the logit transform of loan default rate series, "$\Delta logit$". Since the DY methodology, based on the GFEVD, requires the VAR residuals to be Gaussian, I employ the Jarque-Bera (JB) test. As can be seen from Table 2.5, the null hypothesis, that is the estimated residuals are normally distributed, cannot be rejected at 95% confidence level, for most of the loan default rates series.

Table 2.3: Augmented Dickey Fuller (ADF) test on the logit trasformation of the regional loan default rates series.

Hypothesis	Statistic	Critical values 1%	5%	10%
$H_0 : \beta = 0$	t-based	-3.51	-2.89	-2.58
$H_0 : \alpha = 0$	t-based	3.22	2.54	2.17
$H_0 : \alpha = \beta = 0$	F-based	6.70	4.71	3.86

Regions	Consumer Households				Non Financial Firms				Producer Households			
	$H_0 : \beta = 0$	$H_0 : \alpha = 0$		$H_0 : \alpha = \beta = 0$	$H_0 : \beta = 0$	$H_0 : \alpha = 0$		$H_0 : \alpha = \beta = 0$	$H_0 : \beta = 0$	$H_0 : \alpha = 0$		$H_0 : \alpha = \beta = 0$
	τ	t-value	p-value	Γ	τ	t-value	p-value	Γ	τ	t-value	p-value	Γ
PIEDMONT	-3.017	-3.099	0.003	5.437	-1.065	-1.011	0.316	0.728	-2.215	-2.222	0.030	2.480
AOSTA VALLEY	-3.358	-3.413	0.001	5.950	-2.616	-2.581	0.012	3.501	-4.124	-4.071	0.000	8.536
LIGURIA	-2.521	-2.635	0.010	4.070	-2.620	-2.612	0.011	3.433	-2.653	-2.675	0.009	3.609
LOMBARDY	-2.416	-2.450	0.017	3.093	-1.262	-1.236	0.221	0.825	-1.756	-1.757	0.083	1.544
VENETO	-2.955	-3.004	0.004	4.715	-0.205	-0.122	0.903	0.411	-2.241	-2.247	0.028	2.528
FRIULI VG	-4.546	-4.622	0.000	11.173	-1.096	-1.057	0.294	0.667	-2.856	-2.847	0.006	4.087
EMILIA-ROMAGNA	-2.072	-2.104	0.039	2.353	-1.446	-1.400	0.166	1.115	-1.397	-1.391	0.169	0.981
TRENTINO AA	-1.994	-2.011	0.048	2.066	-2.262	-2.251	0.028	2.561	-3.224	-3.220	0.002	5.197
TUSCANY	-3.706	-3.833	0.000	8.343	-0.491	-0.435	0.665	0.278	-1.257	-1.245	0.217	0.806
UMBRIA	-2.435	-2.470	0.016	3.115	-1.106	-1.068	0.289	0.703	-2.369	-2.372	0.020	2.814
MARCHE	-2.339	-2.376	0.020	2.910	-0.954	-0.858	0.394	0.698	-2.199	-2.192	0.032	2.421
LAZIO	-2.674	-2.898	0.005	5.339	-1.374	-1.355	0.180	0.949	-1.899	-1.994	0.050	2.293
ABRUZZO	-1.477	-1.552	0.125	1.343	-1.874	-1.845	0.069	1.771	-1.973	-2.010	0.048	2.057
MOLISE	-2.054	-2.136	0.036	2.413	-2.122	-2.105	0.039	2.252	-2.992	-3.001	0.004	4.507
CAMPANIA	-2.462	-2.642	0.010	4.186	-1.980	-2.008	0.049	2.031	-2.244	-2.314	0.024	2.840
APULIA	-2.340	-2.493	0.015	3.550	-2.054	-2.093	0.040	2.212	-2.371	-2.443	0.017	3.223
BASILICATA	-1.653	-1.821	0.073	2.092	-2.782	-2.782	0.007	3.880	-2.963	-2.984	0.004	4.482
CALABRIA	-2.564	-2.718	0.008	4.200	-2.742	-2.800	0.007	3.990	-2.971	-2.963	0.004	4.414
SICILY	-1.990	-2.113	0.038	2.607	-1.473	-1.492	0.140	1.120	-1.621	-1.622	0.109	1.317
SARDINIA	-2.329	-2.559	0.013	4.241	-1.282	-1.286	0.203	0.827	-1.449	-1.464	0.148	1.075

Note. The table reports the ADF statistics computed for the logit transformation of the 20 regional series of Default rate on loan facilities for each of the three private sector categories, which run from $1996Q_1$ to $2015Q_4$, considering a model with an intercept and a number of lags equal to 4. The model has the following representation: $\Delta y_t = \alpha + \beta y_{t-1} + \Theta_1 \Delta y_{t-1} + \Theta_2 \Delta y_{t-2} + \Theta_3 \Delta y_{t-3} + \Theta_4 \Delta y_{t-4} + \varepsilon_t$. In particular, three hypotheses are tested: *1)* $H_0 = \beta = 0$, *2)* $H_0 : \alpha = 0$ and *3)* $\alpha = \beta = 0$. The table reports the statistics of interest: τ refers to the hypothesis (1), that is the presence of a unit root in the processes, t-value refers to the hypothesis (2), that is the significance of the intercept (p-values for this statistic are also reported), and Γ refers to the hypothesis (3), that is the joint hypothesis that there is a non-stationary process without drift. Each statistic is compared with the corresponding critical values reported in the upper-right sub-table.

Table 2.4: KPSS unit root test on the logit transformation of the regional default rates series.

	Consumer Households			Non Financial Firms			Producer Households		
	$\ell = 2$	$\ell = 4$	$\ell = 6$	$\ell = 2$	$\ell = 4$	$\ell = 6$	$\ell = 2$	$\ell = 4$	$\ell = 6$
PIEDMONT	0.532	0.351	0.269	0.368	0.264	0.215	0.427	0.306	0.244
AOSTA VALLEY	0.346	0.280	0.240	0.152	0.173	0.168	0.057	0.061	0.072
LIGURIA	0.583	0.382	0.287	0.381	0.271	0.224	0.420	0.319	0.264
LOMBARDY	0.446	0.309	0.245	0.553	0.354	0.271	0.513	0.336	0.259
VENETO	0.491	0.324	0.251	0.578	0.375	0.282	0.403	0.281	0.226
FRIULI VG	0.456	0.319	0.257	0.333	0.255	0.217	0.384	0.297	0.259
EMILIA ROMAGNA	0.529	0.350	0.265	0.341	0.244	0.203	0.477	0.330	0.262
TRENTINO AA	0.402	0.300	0.245	0.352	0.272	0.234	0.135	0.124	0.120
TUSCANY	0.600	0.386	0.294	0.529	0.346	0.264	0.569	0.373	0.286
UMBRIA	0.408	0.318	0.266	0.417	0.322	0.260	0.278	0.206	0.173
MARCHE	0.525	0.345	0.267	0.407	0.289	0.233	0.330	0.243	0.207
LAZIO	0.613	0.401	0.301	0.554	0.367	0.278	0.592	0.384	0.290
ABRUZZO	0.514	0.337	0.257	0.435	0.303	0.243	0.529	0.358	0.274
MOLISE	0.448	0.329	0.266	0.203	0.160	0.135	0.228	0.196	0.169
CAMPANIA	0.611	0.397	0.296	0.575	0.380	0.290	0.608	0.401	0.306
PUGLIA	0.625	0.401	0.301	0.583	0.386	0.294	0.601	0.393	0.299
BASILICATA	0.496	0.343	0.269	0.344	0.273	0.229	0.178	0.211	0.209
CALABRIA	0.601	0.391	0.296	0.554	0.370	0.287	0.345	0.268	0.236
SICILY	0.606	0.391	0.294	0.607	0.387	0.291	0.531	0.367	0.280
SARDINIA	0.595	0.388	0.289	0.609	0.395	0.296	0.564	0.372	0.284

Note. The table reports the KPSS unit root test statistic, computed for the logit transformation of the 20 regional series of loan default rates, where the null hypothesis is that the series is stationary around a trend. The test statistics are reported for each of the three private sector categories and for different lag parameter truncation, $\ell = 2$, $\ell = 4$ and $\ell = 6$ (see Kwiatkowski *et al.*, 1992). The critical value at 5% (0.146) is also reported.

2.5 Results

2.5.1 Total connectedness index

I compute the *Total connectedness* index, e.g. a proxy of the spatial dependence, by taking the average of the off-diagonal elements in the generalized forecast error variance decomposition (GFEVD) matrix. In this analysis, I focus on a forecast horizon equal to four quarters ($H = 4$). The index provides a measure of the total connection between regional default rates on loan facilities (as suggested by Diebold & Yilmaz (2012, 2014), see eq.(2.13)). First, I focus on the static measure of the *Total connectedness* index which is obtained through the estimation of the lasso VAR(1) model fitted to the $K = 20$ regional default rate series over the full sample period ($1996Q2 - 2015Q4$). The estimation exercise and the corresponding results, reported in the rest of the Section, refer to an analysis conducted for each of the three private sub-sectors (consumer households, non-financial firms and producer households), separately. The static, unconditional, analysis

Table 2.5: Jarque-Bera test on the residuals of the VAR model fitted to the Δ logit transformation of the regional loan default rates series.

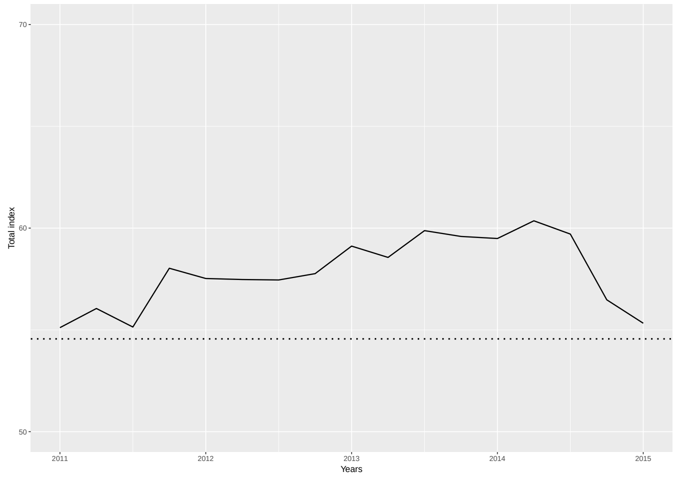
	Consumer Households		Non Financial Firms		Producer Households	
	Statistic	p-value	Statistic	p-value	Statistic	p-value
PIEDMONT	2.498	0.287	4.03	0.133	5.114	0.078
AOSTA VALLEY	1.558	0.459	0.493	0.781	28.634	0.000
LIGURIA	2.785	0.248	1.233	0.540	1.076	0.584
LOMBARDY	480.093	0.000	1.117	0.572	1.592	0.451
VENETO	105.406	0.000	0.206	0.902	1.075	0.584
FRIULI VG	5.101	0.078	3.179	0.204	1.164	0.559
EMILIA ROMAGNA	0.141	0.932	350.322	0.000	13.798	0.001
TRENTINO AA	0.031	0.985	1.584	0.453	6.079	0.048
TUSCANY	7.053	0.029	1.108	0.575	0.466	0.792
UMBRIA	125.254	0.000	0.011	0.994	0.121	0.941
MARCHE	3.091	0.213	1.768	0.413	0.577	0.75
LAZIO	58.893	0.000	0.207	0.902	0.176	0.916
ABRUZZO	4.565	0.102	25.587	0.000	3.246	0.197
MOLISE	0.049	0.976	2.787	0.248	0.275	0.871
CAMPANIA	6.514	0.038	24.836	0.000	1.171	0.557
APULIA	1.890	0.389	8.829	0.012	0.507	0.776
BASILICATA	8.429	0.015	0.672	0.715	4.229	0.121
CALABRIA	4.315	0.116	1.274	0.529	17.018	0.000
SICILY	3.450	0.178	3.247	0.197	1.928	0.381
SARDINIA	0.260	0.878	1.946	0.378	1.843	0.398

Note. The table reports the Jarque-Bera (JB) test statistic computed for the series of residuals obtained through the estimation of a sparse VAR(1) model for the three private sector categories. The statistics are compared with the critical value of a Chi-squared distribution with 2 degrees of freedom, that is $\chi^2(2) = 5.99$, at 5% significance level. P-values are also reported.

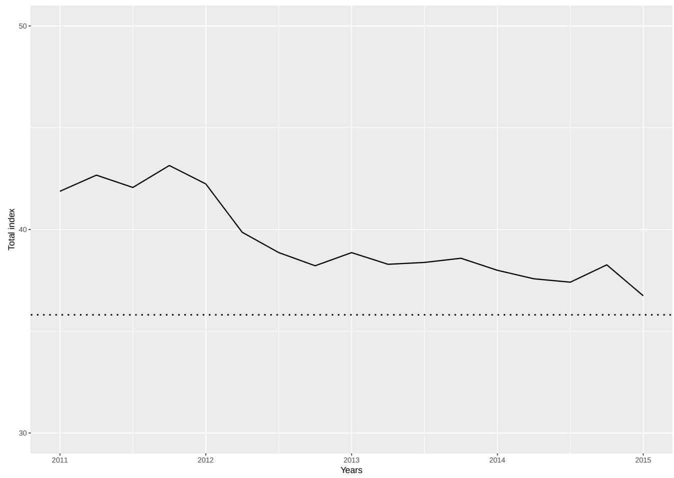
shows that the consumer households sector reports the highest value of the *Total connectedness* measure (54.6%), while the index is relatively lower for producer households (41.3%) and non-financial firms (35.8%).
Following Diebold & Yilmaz (2012, 2014), I also compute a time-varying measure of the *Total connectedness* index as well as of the pairwise and total directional indices suggested by Greenwood-Nimmo *et al.* (2015), using a rolling estimation window width equal to 63 quarters, with a starting sample which covers the $1996Q2 - 2011Q4$ time period. Similarly to the full-sample analysis, I use a forecast horizon equal to four quarters ($H = 4$).

Figure 2.4 shows the time-varying *Total connectedness* index (black line) for consumer households (panel a), non-financial firms (panel b) and producer households (panel c). For each panel, I also report the unconditional values (dotted line) of the index (the ones reported above), which can be interpreted as the long-run equilibrium. As can be seen from Figure 2.4 (panel a), the time-varying analysis shows that there is some evidence of an increase of spatial dependence among consumer households, since the *Total connectedness* index is above the long-run equilibrium over the second part of the

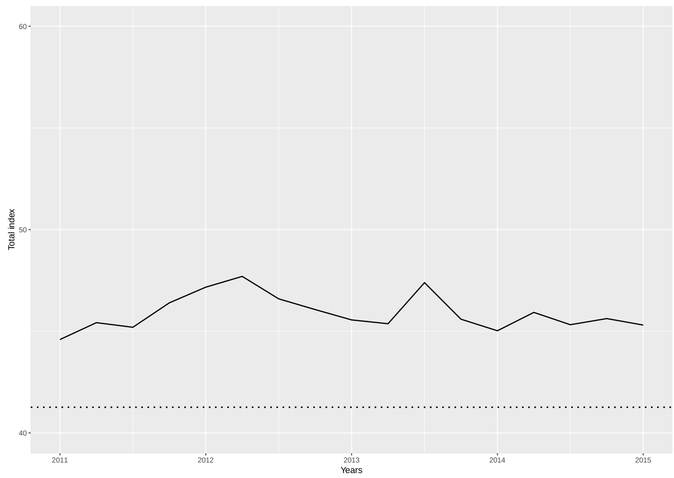
Figure 2.4: Time-varying *Total connectedness* index (in percentage) at $H = 4$ steps ahead, $2011Q4 - 2015Q4$.



(a) Consumers households



(b) Non-financial firms



(c) Producers households

Note. The figure shows the time-varying *Total connectedness* index (black line) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The time-varying *Total connectedness* index is reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c). The static, unconditional, *Total connectedness* index values (dotted line) are also reported: 54.6% (consumer households), 35.8% (non-financial firms) and 41.3% (producer households).

sample. In particular, the *Total connectedness* index rises from 55.1% to 60.4% between 2012Q2 and 2015Q1, before getting back to the long-run value (the index is equal to 55.3% in the last quarter of 2015). Also for producer households (see Figure 2.4, panel c), there is evidence of an increase in the *Total connectedness* index (the average value, over the 2011$Q4$ − 2015$Q4$, is 45.9%) since it is above the long-run value of 41.3%. Given the long-run value of the *Total connectedness* index for non-financial firms equals 35.8%, this private sector category manifests evidence of an increase in the total default connectedness over 2011$Q4$ − 2012$Q4$, and a subsequent fall in the index over the last three years of the sample under investigation (see Figure 2.4, panel b).

2.5.2 GNS Results

In this section, I report the results of the $H = 4$ steps ahead connectedness analysis, conducted by using the approach proposed in the study of Greenwood-Nimmo *et al.* (2015). In particular, these results refer to a specific aggregation scheme of the $K = 20$ regional loan default rate series into $N = 5$ groups of regions, that is Northwest, Northeast, Centre, South and Islands. Similarly to the analysis of the *Total connectedness* index, all the measures reported in this section concern the estimation of the lasso VAR(1) models for each of the three private sub-sectors (consumer households, non-financial firms and producers households).
I first focus on the static connectedness measures obtained by exploiting the full-sample (1996Q2 − 2015Q4) information.

Table 2.6 shows the group connectedness matrix for consumer households (panel a), non-financial firms (panel b) and producer households (panel c). Each panel shows the *Total within-group* forecast error variance (FEV) contributions, for each of the N groups, that is the elements on the main diagonal (see eq.(2.17)), and the off-diagonal elements which measure the pairwise spillovers among the groups (see eqs.(2.18) and (2.19)). It is important to observe that the values reported in Table 2.6, together with all the results presented in the rest of the chapter, are expressed, given the above-mentioned re-normalization proposed by Greenwood-Nimmo *et al.* (2015) (see Section 2.3.2), as a percentage of the FEV computed for the whole system. The *Total within-group* index reflects the importance of the local factors in each group, and the higher is the value associated with this measure the stronger is their contribution to the own-group domestic conditions.

Table 2.6: Group connectedness matrix. Full sample estimation ($1996Q2 - 2015Q4$), $H = 4$ steps ahead.

	Northwest	Northeast	Centre	South	Islands
Northwest	10.786	1.485	2.567	3.817	1.346
Northeast	2.701	12.312	1.665	2.569	0.752
Centre	2.148	1.221	10.642	5.153	0.836
South	3.756	1.867	3.826	18.994	1.557
Islands	1.609	0.535	1.064	2.441	4.350

(a) Consumer households.

	Northwest	Northeast	Centre	South	Islands
Northwest	13.150	1.832	1.065	3.513	0.440
Northeast	2.162	14.341	1.109	1.633	0.755
Centre	2.331	1.690	13.120	2.665	0.194
South	2.292	1.172	1.001	24.717	0.818
Islands	1.206	0.951	0.163	1.737	5.942

(b) Non-financial firms.

	Northwest	Northeast	Centre	South	Islands
Northwest	12.354	1.787	1.986	2.986	0.887
Northeast	2.720	12.816	1.539	2.310	0.615
Centre	2.372	1.493	12.479	3.121	0.535
South	3.167	1.356	1.867	22.812	0.798
Islands	0.530	0.641	0.818	1.101	6.910

(c) Producers households.

Note. The table reports the static group connectedness matrix obtained through a full sample estimation ($1996Q2 - 2015Q4$), by using a forecast horizon equal to four quarters ($H = 4$). The measures are reported for each of the three private sector categories: consumer households (panel a), non-financial firms (panel b) and producer households (panel c). In each panel, the main diagonal elements give the *Total within-group* forecast error variance (FEV) contributions, for each of the $N = 5$ groups (see eq.(2.17)). The off-diagonal elements give the spillover effects among the groups (see eqs.(2.18) and (2.19)). The values are expressed as a percentage of the FEV computed for the whole system.

The results shown in Table 2.6 do not reveal large differences among the three categories of the private sector (consumer households, non-financial firms and producer households). More specifically, in each sector, the *Total within-group* indices tend to be larger than the off-diagonal measures, with the highest values recorded in the South of Italy (19%, 24.7% and 22.8%, respectively). Contrary to the South, the other macro-region in the Mezzogiorno, Insular Italy, shows a relatively small contribution of local factors, for all three private sub-sectors (4.4%, 5.9% and 6.9%). The results for Northwest, Northeast and Centre are similar among the three private sub-sectors. The values of the index for consumer households are 10.8%, 12.3% and 10.6%, respectively; the ones for non-financial firms are 13.2%, 14.3% and 13.1%, respectively, and the ones for producer households are 12.4%, 12.8% and 12.5%, respectively.

2.5.2.1 Static total directional analysis

I also focus on the Dependence score which is presented in Table 2.7, together with the other aggregate connectedness measures. The Dependence score ($\mathcal{O}_n^H$), with $0 \leq \mathcal{O}_n^H \leq 1$, measures the relative importance of an external shock for a certain group. Large values ($\mathcal{O}_n^H \rightarrow 1$) indicates that the group largely depends on external conditions, while small values ($\mathcal{O}_n^H \rightarrow 0$) reveal low degree of exposure to external shocks. The results in Table 2.7 indicate that Insular Italy has the highest dependence value for consumer households (0.57) and non-financial firms (0.41), decisively above the corresponding average values (0.45 and 0.31, respectively), while the scores are more similar for producer households, with the Northwest and Centre of Italy sharing the largest value (0.38). These results are also presented in three quantile maps, one for each private sub-sector (see Figure 2.5, panel a).
Additional information on the transmission mechanism of spillovers among groups might arise from the aggregate measures presented in Table 2.7.
In particular, I focus on those measures which provide information on the role played by a specific group as a shock contributor (or receiver).

The contribution of specific-group conditions to the FEV of the whole system is measured by the *To* index. It can be seen from the results in Table 2.7 that the group contributing the most is the South of Italy, where the values of the *To* index, 13.98% (consumer households), 9.55% (non-financial firms) and 9.52% (producer households), decisively exceed the corresponding average values (8.58%, 5.75% and 6.53%, respectively). The next largest values are reported by the Northwest of Italy: 10.21% (consumer households), 7.99% (non-financial firms) and 8.79% (producer households). Oppositely, I find that Insular Italy has the lowest contribution to the whole FEV for consumer households (4.50%), non-financial firms (2.21%) and producer households (2.84%), less than half of the average values.

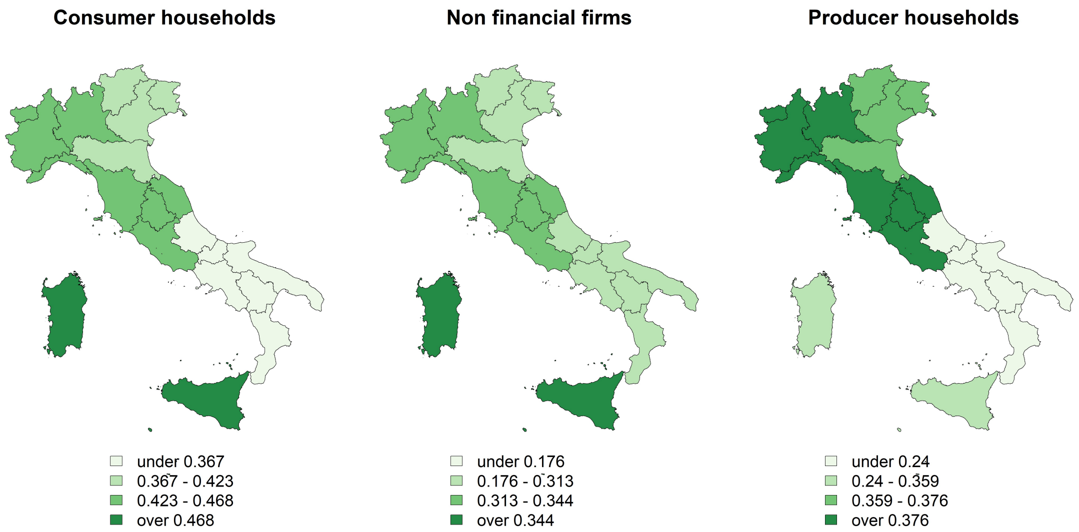
These results are confirmed by looking at the *Net* index shown in Table 2.7. In fact, the net contributor in terms of spillovers is the South of Italy in all the three private sub-sectors: 2.98% (consumer households), 4.27% (non-financial firms) and 2.33% (producer households). These high values, driven by the remarkable relative large magnitude of the *TO* indices, highlight a leading role of the South of Italy in contributing to the system-wide risk. The second ranked is the Northwest of Italy: 1.00% (consumer households), 1.14% (non-financial firms) and 1.14% (producer households). If I focus on the lowest values reported in Table 2.7, the ranking reveals that the net receiver is the Northeast of Italy for consumer households and producer households, −2.58% and −1.91% respectively, while the smallest *Net* index is reported by the Centre of Italy, −3.54%,

Table 2.7: Aggregate connectedness measures. Full sample estimation ($1996Q2 - 2015Q4$), $H = 4$ steps ahead.

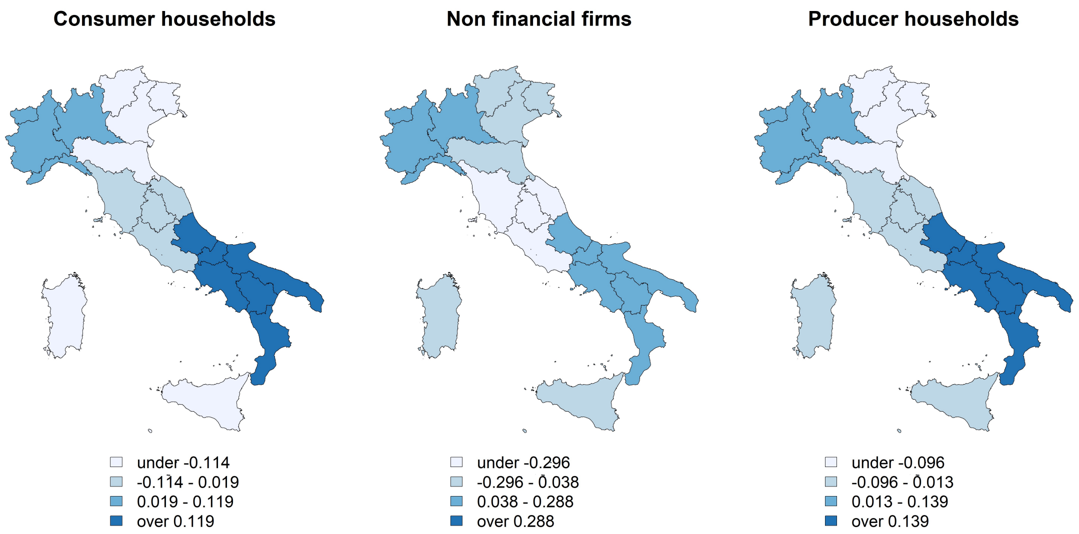
	Within	From	To	Net	Dep.	Infl.
Consumers						
Northwest	10.786	9.214	10.214	1.000	0.461	0.051
Northeast	12.312	7.688	5.108	-2.580	0.384	-0.202
Centre	10.642	9.358	9.122	-0.236	0.468	-0.013
South	18.994	11.006	13.980	2.974	0.367	0.119
Insular	4.350	5.650	4.491	-1.158	0.565	-0.114
Average	11.417	8.583	8.583	0.000	0.449	-0.032
Firms						
Northwest	13.150	6.850	7.991	1.141	0.343	0.077
Northeast	14.341	5.659	5.645	-0.014	0.283	-0.001
Centre	13.120	6.880	3.339	-3.541	0.344	-0.347
South	24.717	5.283	9.549	4.266	0.176	0.288
Insular	5.942	4.058	2.206	-1.852	0.406	-0.296
Average	14.254	5.746	5.746	0.000	0.310	-0.056
Producers						
Northwest	12.354	7.646	8.789	1.143	0.382	0.070
Northeast	12.816	7.184	5.277	-1.907	0.359	-0.153
Centre	12.479	7.521	6.210	-1.311	0.376	-0.096
South	22.812	7.188	9.518	2.330	0.240	0.139
Insular	6.910	3.090	2.835	-0.255	0.309	-0.043
Average	13.474	6.526	6.526	0.000	0.333	-0.017

Note. The table reports the values of the Within, From, To and Net measures computed according to eqs.(2.17) and (2.20), for each of the three private sector categories: consumer households, non-financial firms and producer households. The values of these four indices are expressed as a percentage of the FEV computed for the whole system. Dep. denotes the dependence index, O_n^H, $0 \leq O_n^H \leq 1$ (see eq.(2.21)), while Infl. denotes the influence index, I_n^H (see eq.(2.22)).

Figure 2.5: Dependence and Influence indices. Full sample estimation ($1996Q2 - 2015Q4$), $H = 4$ steps ahead. $N = 5$ Italian groups of regions.



(a) Dependence index quantile maps.



(b) Influence index quantile maps.

for non-financial firms.

Finally, the *Influence* index ($-1 \leq \mathcal{I}_n^H \leq 1$) provides a measure of the role played by a specific group as net receiver ($-1 \leq \mathcal{I}_n^H < 0$), transmitter ($0 < \mathcal{I}_n^H \leq 1$), or neither a net receiver or transmitter ($\mathcal{I}_n^H = 0$). Substantially, for each group this score is computed as the *Net* index normalized by the sum between the *From* and *To* measures. Therefore, the results in Table 2.7, together with the quantile maps shown in Figure 2.5 (panel b), display additional evidence of the bigger role played by the South of Italy as net influencer. In fact, for all the private sub-sectors, Southern regions show high values of the index (0.12, 0.29 and 0.14, respectively), decisively above the corresponding average values. Positive values are also reported by the Northwest of Italy: 0.05 (consumer households), 0.08 (non-financial firms) and 0.07 (producer households). The remaining groups of regions report negative values of the *Influence* score. The Northeast of Italy presents large negative values of the score for consumer households (-0.20) and producer households (-0.15), while Central Italy (-0.35) and Insular Italy (-0.30) are the largest net shock recipients for non-financial firms.

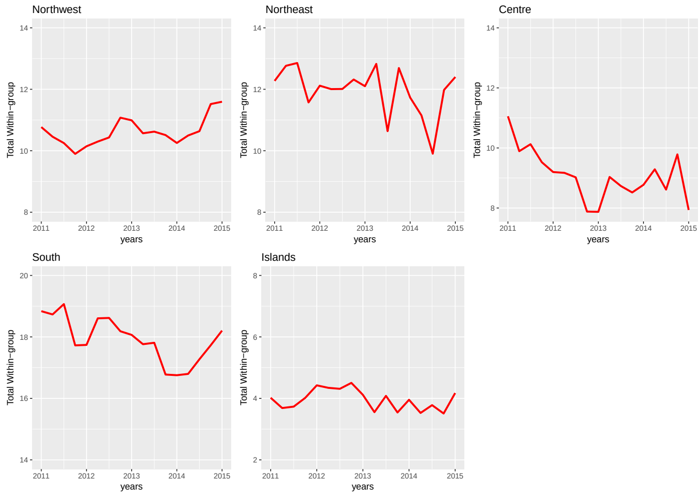
2.5.2.2 Rolling total directional analysis

Figures 2.6-2.10 show the time-varying connectedness measures.

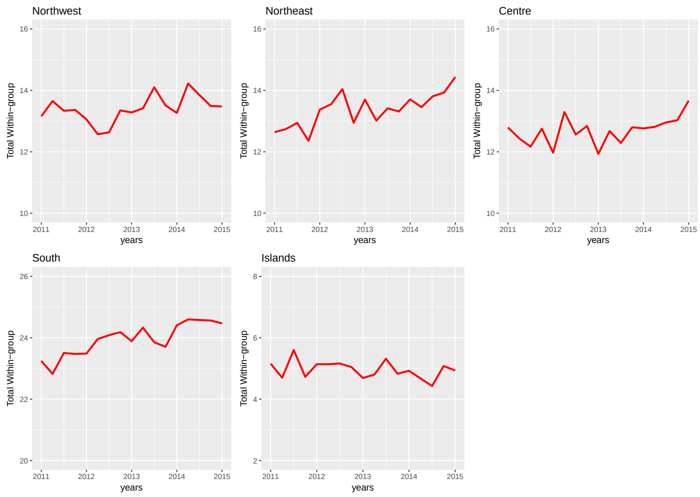
Figure 2.6 shows the time-varying total *Within-group* index for all the three private sub-sectors. From the chart, it can be seen that the results obtained for the unconditional analysis are also valid in the dynamic scenario. In fact, the values of the index reported by consumer households tend to be smaller than the ones showed by non-financial firms and producer households. Furthermore, notwithstanding a reduction reported for consumer households and producer households, the South of Italy shows the highest *Within-group* for all the three private sub-sectors, during the entire sample period.

Figure 2.6 (panel a), which reports the results for consumer households, highlights a decreasing trend in the *Within-group* index in South of Italy (from 18.84% to 16.75%), with the exception of the last 3 quarters, when the index increases again reaching 18.20%, and Central Italy (-3.12% is the overall reduction during the whole sample). In the Northwest of Italy, the index is stable around 10.50% until 2014$Q4$, before increasing by 1.10 point percentage in the subsequent 4 quarters. The results for non-financial firms shown in Figure 2.6 (panel b) reveal an overall increase in the own-group measure reported by the South of Italy (from 23.24% to 24.47%) and the Northeast of Italy (from 12.64% to 14.43%). For producer households, I find that the index increases in

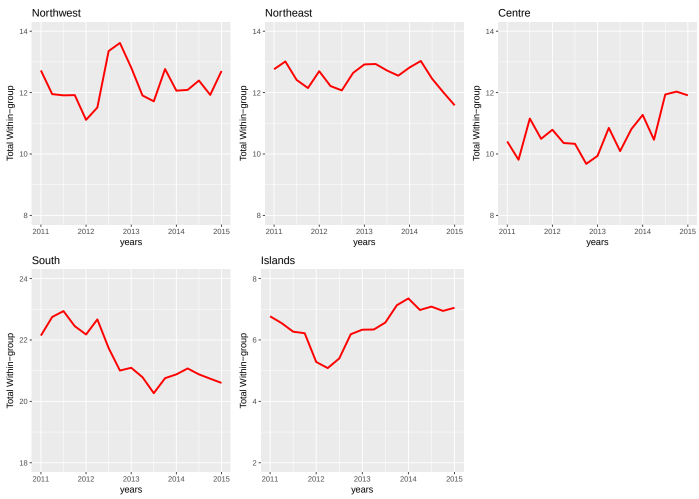
Figure 2.6: Time-varying total *Within-group* index at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. $N = 5$ Italian groups of regions.



(a) Consumer households



(b) Non-financial firms



(c) Producer households

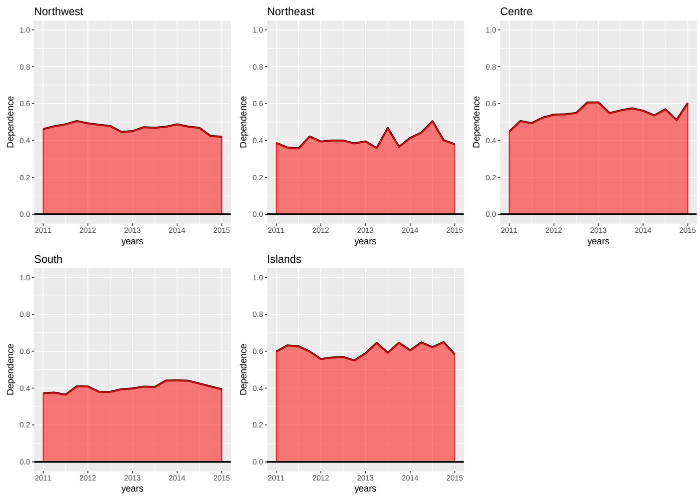
Note. The figure shows the time-varying total *Within-group* index (see eq.(2.17)) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The index is reported for each of the three private subsectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c). The values are expressed as a percentage of the FEV computed for the whole system.

Central Italy by more than 1.50% during the whole period (10.41% is the value reported in 2011Q4). Moreover, there is evidence of a reduction of the *Within-group* index in Insular Italy until 2013Q1 before increasing in the rest of the sample (see Figure 2.6, panel c). Oppositely, as shown in Figure 2.6 (panel c), the value of the index falls in South of Italy, from 22.14% to 20.60% between 2011Q4 and 2015Q4.

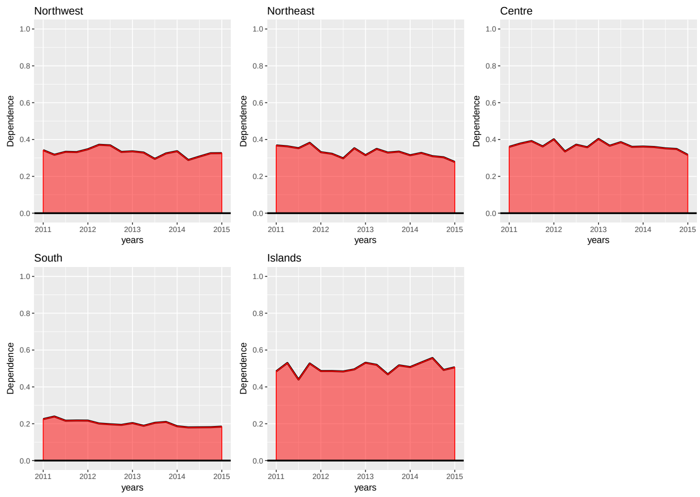
Figure 2.7 displays the time-varying *Dependence* score for consumer households (panel a), non-financial firms (panel b) and producer households (panel c). Similarly to the results obtained from the unconditional analysis, if I now turn to the dynamic analysis the results highlight the large level of dependence reported by Insular Italy (around 0.60), during the whole sample period, together with Central Italy (the index increases from 0.45 to 0.60 since 2011Q4, see Figure 2.7, panel a). For non-financial firms, the results show that Insular Italy is the group reporting the largest degree of dependence from the system, with an average value equal to 0.50 for the entire period (see Figure 2.7, panel b). The results in Figure 2.7 (panel c) show that for producer households the group which reports the highest score is Central Italy, with values of the index ranging from 0.40 to 0.50. High *Dependence* scores are also reported by the Northwest and Northeast of Italy, with the same average value reported during the whole sample (0.38). Finally, the Insular Italy shows an increasing trend in the *Dependence index* over the period 2011$Q4 - 2013Q$1 reaching its peak (0.49), before reducing to 0.30 at the end of the sample period.

Figure 2.8 presents the *To* connectedness index obtained from the rolling-window estimation. The charts shown in Figure 2.8 (panel a) validate the unconditional results, that is a relevant contribution to the system-wide FEV arising from the South of Italy during the whole period (14% on average), for consumer households. In Central Italy, the index is relatively stable around $8 - 9$%, before falling in the last 2 quarters of the sample (from 9.02% to 6.19%). Oppositely, Insular Italy shows a marked increase by 3.37% in the value of the index (the value is 4.71% at the begin of the sample period). For non-financial firms, Figure 2.8 (panel b) highlights relative low values of the *To* index, in particular those reported by Central Italy (4.07%, on average) and Insular Italy (2.56%, on average), together with a sharp decline reported by the South of Italy since 2012Q3 (from 12.37% to 8.82%). Similarly, the results for producer households (Figure 2.8, panel c) show that the index falls by 3.84 point percentage in South of Italy, after reaching a peak in 2013Q1, while there is evidence of a relevant increase in Northwest of Italy during the 2014$Q1 - 2015Q$4 time span, when the index reaches its maximum value (11.87%).

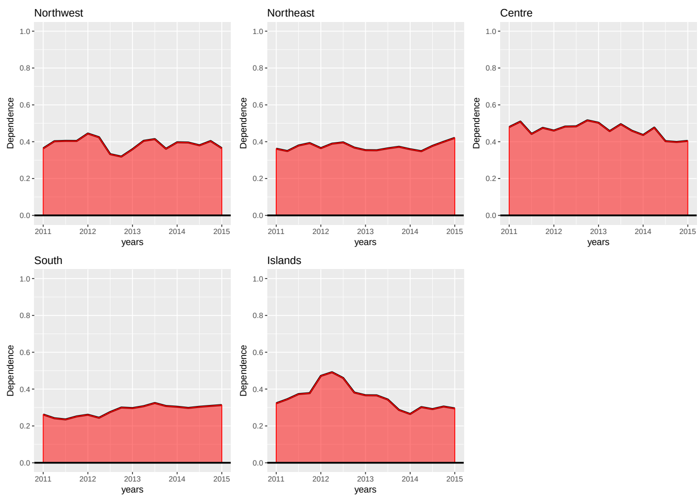
Figure 2.7: Time-varying Dependence index at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. $N = 5$ Italian groups of regions.



(a) Consumer households



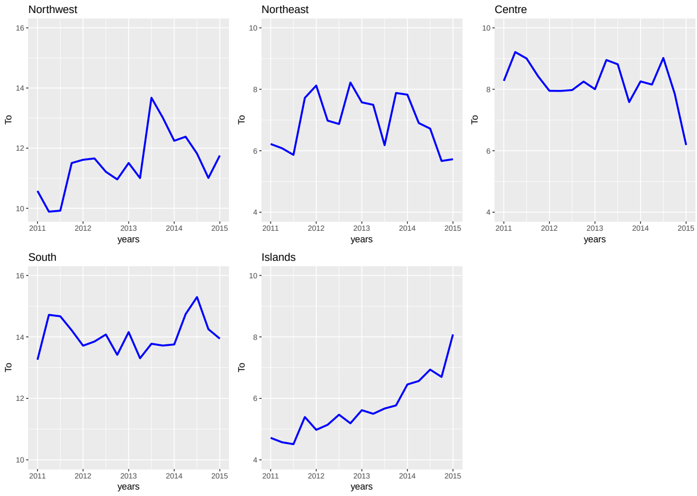
(b) Non-financial firms



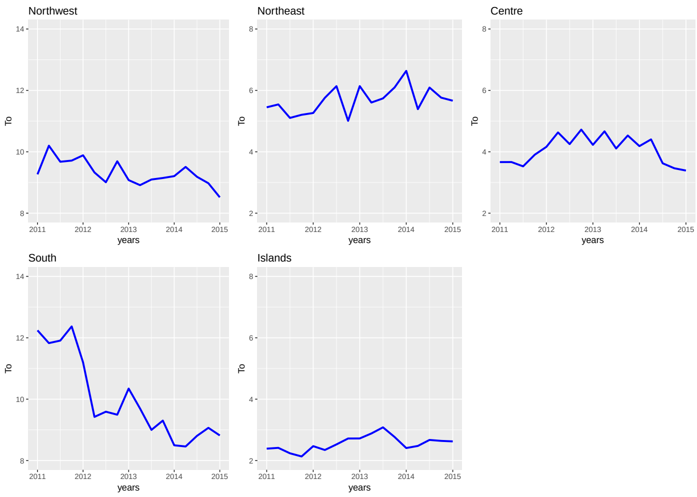
(c) Producer households

Note. The figure shows the time-varying Dependence index (see eq.(2.21)) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The index is reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c).

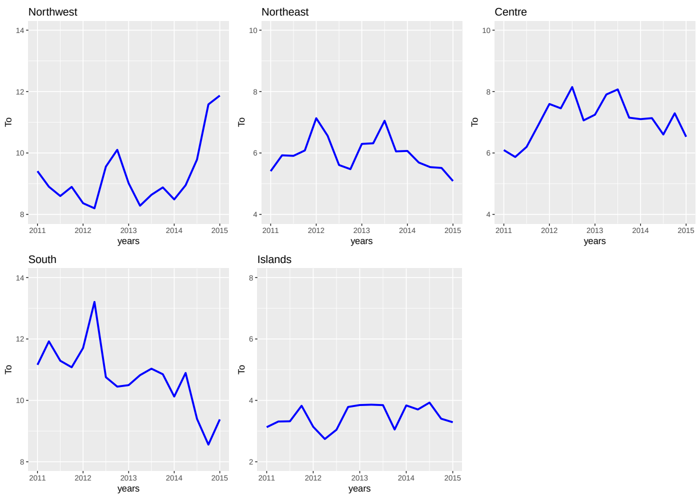
Figure 2.8: Time-varying To index at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. $N = 5$ Italian groups of regions.



(a) Consumer households



(b) Non-financial firms



(c) Producer households

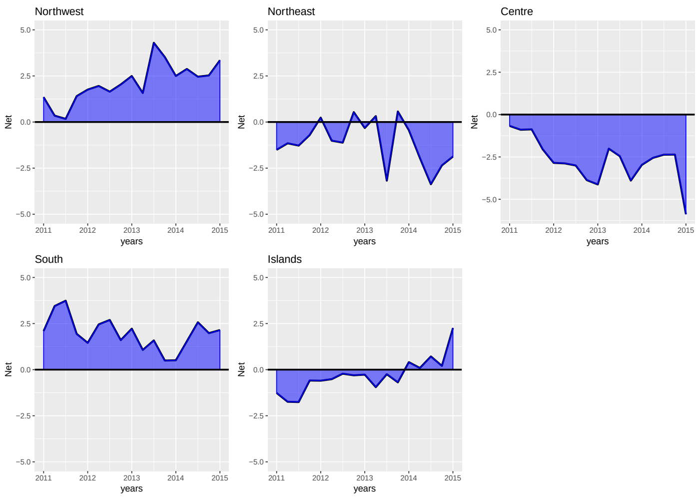
Note. The figure shows the time-varying To connectedness index (see eq.(2.20)) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The index is reported for each of the three private subsectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c). The values are expressed as a percentage of the FEV computed for the whole system.

Figures 2.9 and 2.10 show the results for the time-varying *Net* and *Influence* index. As mentioned before, these connectedness measures provide information on the role played by a group (or entity) as net shocks transmitter or receiver. The *Net* and *Influence* index are similar by construction (see Figures 2.9 and 2.10). In fact, the *Influence* score for the i-th group is the ratio of its *Net* index to the importance of spillovers for that group (measured by the sum of its *From* and *To* index). This normalization allows to obtain values ranging from -1 to 1.

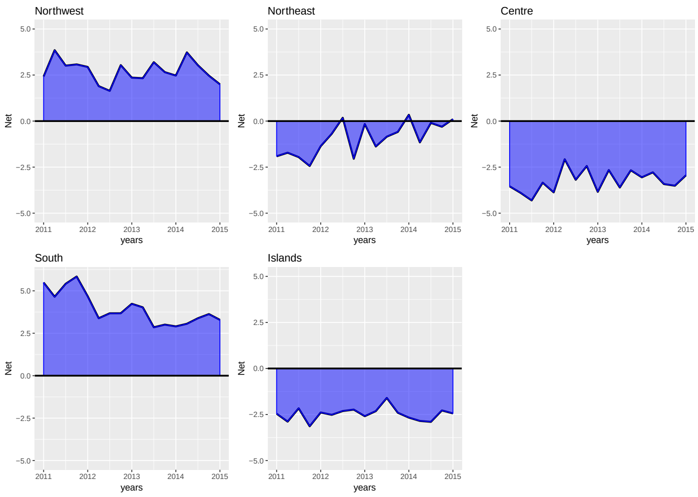
For these reasons, let me focus on commenting the results shown in Figure 2.10. The predominant role played by South of Italy in the static analysis is not confirmed in the dynamic estimation, for all the three private sub-sectors. For example, for consumer households there is evidence of an increasing trend of the *Influence* score reported by the Northwest of Italy since $2012Q2$ (from around zero to 0.17 in $2015Q4$), reaching values of the index higher than the ones presented by Southern regions (see Figure 2.10, panel a). Similarly, the relevant role played by South of Italy sharply decreases for producer households. In fact, as shown in Figure 2.10 (panel c), in spite of relative large values reported in the first part of the sample (with values of the score ranging around 0.20), the *Influence* score declines, reaching negative values in the last two quarters. For non-financial firms, as can be seen from Figure 2.10 (panel b), the South of Italy presents the highest *Influence* score (0.24, on average), together with the Northwest of Italy (0.17, on average). Oppositely, Central Italy presents large negative values of the index for all the three private sub-sectors during the whole time span. However, closer inspection of the charts show that for consumer households and producer households, also Insular Italy plays a negative role as net influencer, at least in the first part of the sample, say since $2013 - 2014$, before becoming positive in the last few quarters (see Figure 2.10, panel a and panel c). For non-financial firms, Insular Italy shows negative values, sharing the role of the group most influenced by the system together with Central Italy (see Figure 2.10, panel b).

The total directional indices provide aggregate information on dependence (influence) of one macro-region from (to) the rest of the country. Since my aim is to detect spatial dependence arising from an increase in the "functional" distance (due to the Consolidation process involving the Italian banking system), I focus on pairwise spillover analysis. The full sample (static analysis) will explore all pairwise effects and the dynamic analysis based on rolling regression will focus only on the effects between Northern and Mezzogiorno regions.

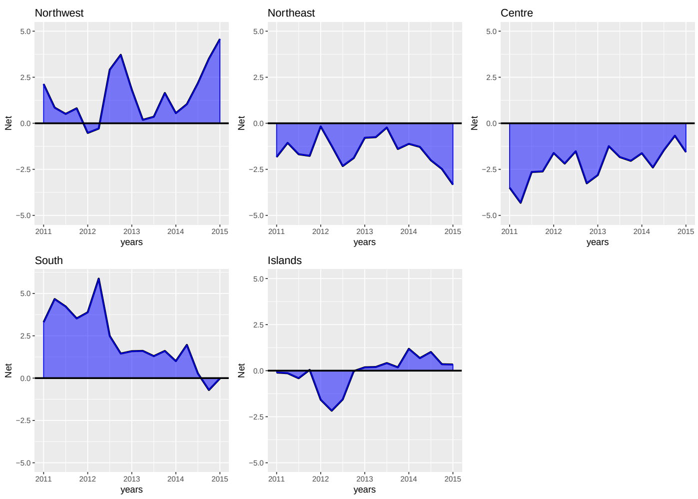
Figure 2.9: Time-varying *Net* index at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. $N = 5$ Italian groups of regions.



(a) Consumer households



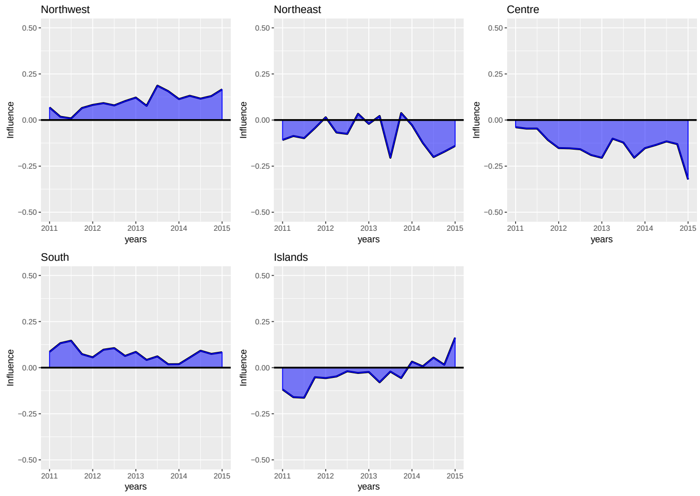
(b) Non-financial firms



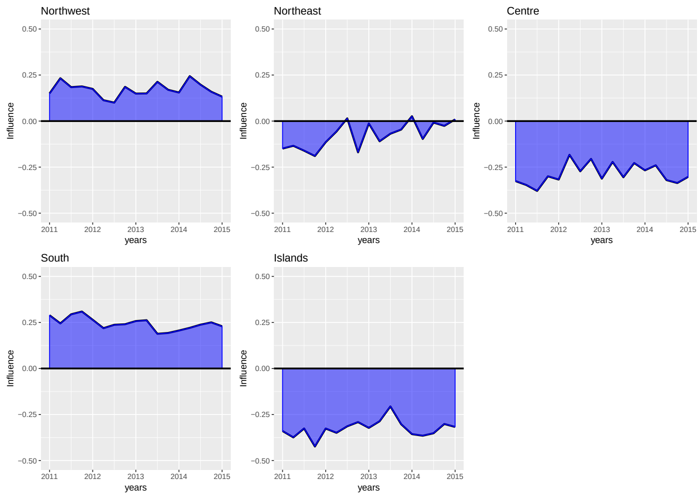
(c) Producer households

Note. The figure shows the time-varying *Net* index (see eq.(2.20)) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2-2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The index is reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c). The values are expressed as a percentage of the FEV computed for the whole system.

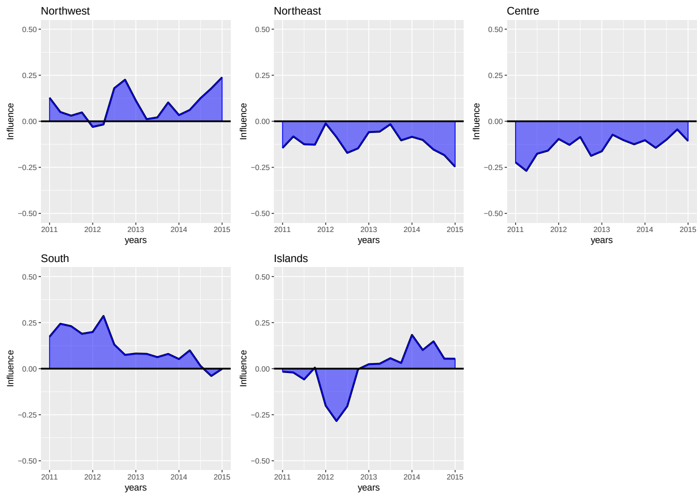
Figure 2.10: Time-varying *Influence* index at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. $N = 5$ Italian groups of regions.



(a) Consumer households



(b) Non-financial firms



(c) Producer households

Note. The figure shows the time-varying *Influence* index (see eq.(2.22)) using a rolling estimation window width equal to 63 quarters, with a starting sample which covers $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). The index is reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c).

Table 2.8: Relative (to *Within-group* index) group connectedness matrix. Full sample estimation ($1996Q2 - 2015Q4$), $H = 4$ steps ahead.

	Northwest	Northeast	Centre	South	Islands
Northwest	100.000	13.766	23.798	35.385	12.480
Northeast	21.940	100.000	13.521	20.866	6.111
Centre	20.182	11.471	100.000	48.420	7.858
South	19.774	9.829	20.143	100.000	8.195
Islands	36.983	12.307	24.464	56.119	100.000

(a) Consumer households.

	Northwest	Northeast	Centre	South	Islands
Northwest	100.000	13.931	8.101	26.717	3.346
Northeast	15.077	100.000	7.733	11.388	5.263
Centre	17.764	12.879	100.000	20.313	1.479
South	9.273	4.742	4.051	100.000	3.308
Islands	20.305	16.011	2.741	29.242	100.000

(b) Non financial firms.

	Northwest	Northeast	Centre	South	Islands
Northwest	100.000	14.464	16.075	24.167	7.184
Northeast	21.223	100.000	12.011	18.024	4.797
Centre	19.010	11.968	100.000	25.010	4.286
South	13.883	5.946	8.184	100.000	3.496
Islands	7.667	9.270	11.835	15.936	100.000

(c) Producers households.

Note. The table reports the static group connectedness matrix obtained through a full sample estimation ($1996Q2 - 2015Q4$), by using a forecast horizon equal to four quarters ($H = 4$). This table is constructed through a re-normalization of Table 2.6. In particular, the (i, j)-th element entering each panel is normalized with respect to the *Total within-group* index of group i. The measures (in percentage) are reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c).

2.5.2.3 Pairwise static analysis

I investigate the pairwise spillovers between groups, that is the measures entering in the off-diagonal elements of the group connectedness matrix (see Table 2.6). Each of those elements measures the contribution to the FEV of group i arising from group j (see eqs.(2.18) and (2.19)).

Since the group connectedness matrix is not row-standardized, to better compare the single contribution of a certain group to the FEV of the others, I need to compute spillover measures that are normalized with respect, for example, to the importance of each within-group condition. To this end, I compute the ratio between the contribution of group j to the FEV of group i and the total *Within-group* index reported by group i (see Table 2.8).

In general, the pairwise spillover analysis shows a large contribution of the Southern regions to the FEV of the other macro-regions, with the exception of the Northest which is more affected by the Northwest. Whilst there is evidence of spatial spillover from South to Northern regions, this is not true for the Islands, which are strongly affected by Northern regions, especially for consumer households and non-financial firms.

As for consumer households (see Table 2.8, panel a), Southern regions show a large contribution to FEVs of other groups. For example, the spillover from South to Northwest accounts for 35.4% of the importance of local factors in Northwest, while the contribution from Northwest to Southern regions is 19.8%. Large values are reported also from South to Centre (48.4%) and from South to Islands (56.1%). The only exception is the largest contribution from Northwest to Northeast (21.9%), slightly above than the spillover that Northeast receives from South (20.9%). Central Italy reports large values of the cross-group measures, including: the contribution to the FEV of Northwest (23.8%) and the one of South (20.1%). The Islands are largely affected by Northern regions (especially from the Northwest), with spillover indices equal to 37% (from Northwest) and 12.3% (from Northeast). Oppositely, there is no evidence of default spillovers from the Islands to Northern regions.

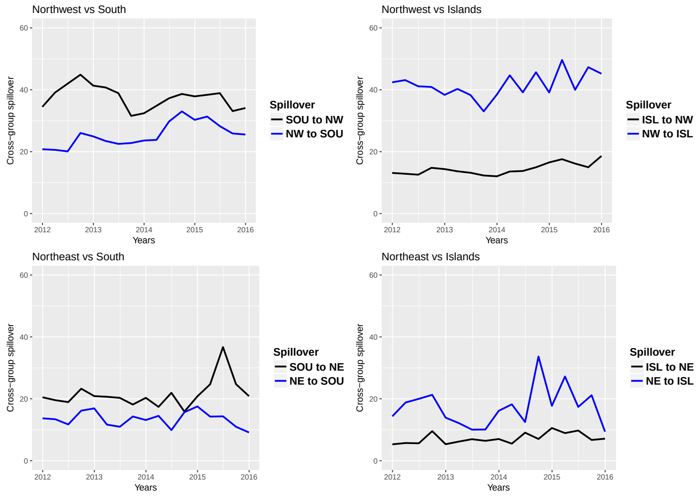
The largest contribution from Northwest to Northeast is more evident looking at the results for non-financial firms (see Table 2.8, panel b). In fact, the value of the cross-group measure is equal to 15.1%, decisively larger than the spillover from South to Northeast (11.4%). Focusing on the other pairwise measures, there is evidence of a large contribution from South to Northwest, 26.7% (the spillover from Northwest to South accounts for less than 1/10 of its within-group measure), and from South to Islands, 29.2% (the spillover from Islands to South is only equal to 3.3%). The Centre of Italy largely receives from both South (20.3%) and Northwest (17.8%), while its contribution to the FEVs of other groups is negligible. Similarly to the results obtained for consumer households, there is a large spillover effect from Northwest to Islands, 20.3% (the index measuring the spillover from Islands to Northwest is only 3.3%), while the spillover from the Northeast to Islands is lower, 16% (still above the spillover arising from Islands to Northeast).

Finally, the results corresponding to pairwise spillover for producer households (see Table 2.8, panel c) are similar to the results for non-financial firms. More specifically, there is evidence of a large contribution from Northwest to Northeast, 21.2% (the spillover from Northeast to Northwest is 14.5%). Once again, Southern regions show the largest pairwise contributions, including: the one to the FEV of the Northwest (24.2%), the Central Italy (25%) and the Islands (15.9%). The Islands are affected, also, by shocks arising from Central Italy (the value of the spillover is 11.8%) and, to less extent, from Northeast (9.3%) and Northwest (7.7%). Large spillover effects are also from Northwest to Central Italy (19%).

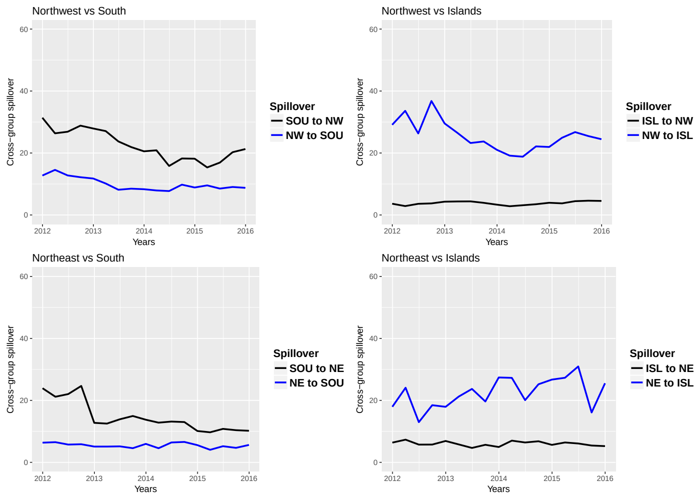
2.5.2.4 Pairwise rolling analysis

The static results are confirmed by using the time-varying cross-group spillovers computed over the $2011Q4 - 2015Q4$ period, with a forecast horizon equal to four quarters ($H = 4$) (see Figure 2.11).
As for consumer households (see Figure 2.11, panel a), the dynamic spillover index from South to Northern regions is permanently above than the one from North to South, over the whole forecast period. In particular, the average value of the spillover from South to Northwest is 37.6% (slightly above the long-run value which is equal to 35%), while the spillover from South to Northeast is 21.5% (the corresponding long-run value is 20.9%). Oppositely, both Northwest and Northeast show a dynamic spillover effect to the Islands larger than the one measured from the Islands to the Northern regions. The difference between the spillover effects is particularly evident looking at the dynamic cross-group measure from Northwest to Islands, whose average value is equal to 41.6% (above the static pairwise measure, 37%).
The large contribution from South to Northern regions is confirmed also for non-financial firms (see Figure 2.11, panel b). However, whilst the spillover from South to Northwest is above its long-run value (26.7%) over the period $2011Q4-2013Q1$, the dynamic spillover decreases since the next quarter, showing an average value equal to 19.4%, over the rest of the forecast period. The spillover from Northwest to South is in line with its long-run value (9.3%). A similar pattern is found in the causality relationship between Northeast and South-Italy. In particular, the spillover from South to Northeast decreases from 24% to 10.2% over the entire forecast period (the long-run value is equal to 11.4%). The spillover from Northern regions to Islands is confirmed, given that I find an average (over the whole forecast period) spillover from Northwest to Islands equal to 25.5% (the long rung value is 20.3%), while the average spillover from Northeast is 22.5% (largely above its long-run value, 16%).
Different results on the comparison between Northwest and South arise from the analysis conducted on the producer households (see Figure 2.11, panel c). The Southern regions show a large spillover to the Northeast over the whole forecast period (with an average dynamic spillover, 20.2%, in line with the corresponding long-run value). The spillovers from South to Northwest decreases over the forecast period. In fact, after increasing in the $2011Q4-2013Q1$, the value of the spillover from South to Northwest shows an average value of 21.7% (lower than the long-run value, 24.2%). The spillover from Northwest to South reports an increase from 16% to 24.4% over the $2011Q4 - 2015Q4$. I also find a decrease in the spillovers from the Northeast to the Islands, especially in the second part of the sample. The spillover from Northeast to Islands shows a large increase in the $2011Q4 - 2013Q1$, before converging to similar values of the spillover arising from the Islands. Finally, the comparison between the dynamic spillovers computed for the

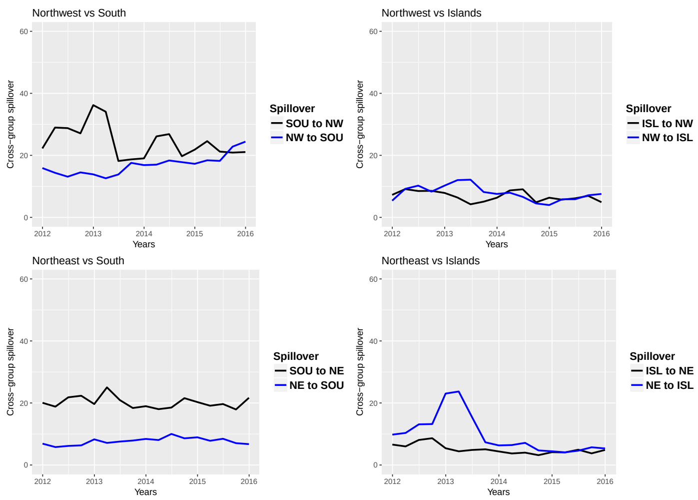
Figure 2.11: Time-varying Cross-group spillovers at $H = 4$ steps ahead, $2011Q4 - 2015Q4$. North vs Mezzogiorno.



(a) Consumer households



(b) Non-financial firms



(c) Producer households

Note. The figure shows the time-varying cross-group spillovers reported in Table 2.8, using a rolling estimation window width equal to 63 quarters, with a starting sample observed over $1996Q2 - 2011Q4$, and a forecast horizon equal to four quarters ($H = 4$). In particular, the figure shows the pairwise spillovers between the Northern regions (Northwest and Northeast) and the Mezzogiorno regions (South and Islands). The measures (in percentage) are reported for each of the three private sub-sectors: consumer households (panel a), non-financial firms (panel b) and producer households (panel c).

Northwest and the Islands does not reveal any additional information with respect to the full sample analysis. In fact, both of the two spillover measures are similar, reporting values in line with the corresponding long-run equilibria.

To summarize, I find evidence of an increase in default rates spatial dependence (relative to the long-run value) for Italian regional default rates over a crisis period ($2011Q4 - 2015Q4$) associated with the last part of the observed sample. These empirical findings are observed for all the three private sector categories, especially for the producer households.

Furthermore, the aggregated total directional indices suggest different dynamics for the two macro-regions of the Mezzogiorno. While the *Influence* index suggests that the South is the largest contributor of shocks to the other macro-regions, Insular Italy shows the highest degree of dependence from the rest of the system (this is particularly true for consumer households and non-financial firms). As for the Northern regions, the *Influence* index suggests that the Northwest is among the largest contributor of shocks to the other macro-regions and the Northeast shows a degree of dependence from the rest of the system similar to Insular Italy.

Furthermore, the comparison of pairwise indices sheds further light on the issue of increasing vulnerability of the Mezzogiorno from the North of Italy as a consequence of the bank consolidation process. In particular, the hypothesis of a bank consolidation process detrimental for the Mezzogiorno is partially supported by the dependence of only Insular Italy on the Northern regions. Moreover, I find evidence of large spillover from the South to the Northwest and Northeast macro-regions.

2.6 Conclusions

In this chapter, I have investigated the spatial spillover effects among 20 Italian regions, by using loans default rates series for consumer households, non-financial firms and producer households, over the $1996Q1 - 2015Q4$ time span. In particular, I use the Diebold-Yilmaz methodology, DY, based on the generalized forecast error variance decomposition (GFEVD) obtained from the estimation, through the Adaptive Elastic net, of a large VAR model, to retrieve a measure of total spatial connectedness among the 20 Italian regional default rates series. Furthermore, the GNS approach enables to compute indices of directional connectedness and, in particular, to assess whether the Mezzogiorno regions are more dependent (relative to the Northern regions) on shocks arising from the other regions.

Using the DY approach to compute an index of total connectedness, the empirical evidence shows an increase in spatial dependence (over the $2011Q4 - 2015Q4$ period) relative to its long-run value. In particular, these empirical findings are more striking for producer households.

I have also focused on indices of directional causality. In this respect, my work is along the lines of Imai & Takarabe (2011) and of Presbitero *et al.* (2014) since the focus is on the role played by large national banks in spreading the crisis from one region to the others within the same country. More specifically, using the GNS approach, I find that Northwest and South are the largest donor of financial stress. These findings, coupled with the analysis of pairwise aggregate spillover effect, partially support the hypothesis of a core-periphery divide and, in particular, the hypothesis of the Mezzogiorno's dependence from the North, triggered by the geographic expansion of Northern banks. This might be motivated by the evidence of large spillovers (both for static and dynamic analysis) from Northern regions only to Insular Italy.

Bibliography

Acharya, Viral V, Hasan, Iftekhar, & Saunders, Anthony. 2006. Should banks be diversified? Evidence from individual bank loan portfolios. *The Journal of Business*, **79**(3), 1355–1412.

Alessandrini, Pietro, Presbitero, Andrea F., & Zazzaro, Alberto. 2009. Banks, distances and firms' financing constraints. *Review of Finance*, **13**(2), 261–307.

Becchetti, Leonardo, Ciciretti, Rocco, & Paolantonio, Adriana. 2014. *Is there a cooperative bank difference?* Working paper series 0314. The Rimini Centre for Economic Analysis.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., & Stein, J. C. 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial economics*, **76**(2), 237–269.

Berger, Allen N, & Udell, Gregory F. 2002. Small business credit availability and relationship lending: The importance of bank organisational structure. *The economic journal*, **112**(477), F32–F53.

Berger, Allen N, & Udell, Gregory F. 2006. A more complete conceptual framework for SME finance. *Journal of Banking & Finance*, **30**(11), 2945–2966.

Bofondi, Marcello, & Gobbi, Giorgio. 2004. *Bad loans and entry into local credit markets.* Temi di discussione (Working papers) No. 509. Bank of Italy.

Bonaccorsi di Patti, Emilia, Eramo, Ginette, & Gobbi, Giorgio. 2005. Piccole e grandi banche nel mercato del credito in Italia. *Banca Impresa Societá, 1*, 3–34.

Bonaldi, Pietro, Hortaçsu, Ali, & Kastl, Jakub. 2015. *An empirical analysis of funding costs spillovers in the euro-zone with application to systemic risk.* NBER Working Paper Series No. 21462. Cambridge, MA: National Bureau of Economic Research.

Brickley, James A, Linck, James S, & Smith, Clifford W. 2003. Boundaries of the firm: evidence from the banking industry. *Journal of Financial Economics*, **70**(3), 351–383.

Cetorelli, Nicola, & Goldberg, Linda S. 2011. Global banks and international shock transmission: Evidence from the crisis. *IMF Economic Review*, **59**(1), 41–76.

Cipollini, Andrea, Lo Cascio, Iolanda, & Muzzioli, Silvia. 2015. *Financial connectedness among European volatility risk premia.* CEFIN Working Papers No 58. University of Modena and Reggio Emilia.

Davis, Richard A, Zang, Pengfei, & Zheng, Tian. 2016. Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics*, **25**(4), 1077–1096.

De Haas, Ralph, & Van Horen, Neeltje. 2011. *Running for the exit: international banks and crisis transmission.* Working paper No. 124. European Bank for Reconstruction and Development.

Demirer, Mert, Diebold, Francis X., Liu, Laura, & Yilmaz, Kamil. 2017. Estimating global bank network connectedness. *Journal of Applied Econometrics*, Online Version of Record published before inclusion in an issue, doi: 10.1002/jae.2585.

Demma, Cristina. 2015. *Localismo bancario e crisi finanziaria.* Questioni di Economia e Finanza (Occasional paper) No. 264. Bank of Italy.

Diebold, Francis X, & Yilmaz, Kamil. 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, **28**(1), 57–66.

Diebold, Francis X, & Yilmaz, Kamil. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, **182**(1), 119–134.

Fan, Jianqing, & Li, Runze. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, **96**(456), 1348–1360.

Foglia, Antonella, Fiori, Roberta, & Iannotti, Simonetta. 2009. *Beyond macroeconomic risk: the role of contagion in the Italian corporate default correlation.* Research Paper No. 12. CAREFIN Research.

Friedman, Jerome, Hastie, Trevor, & Tibshirani, Rob. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, **33**(1).

Giannetti, Mariassunta, & Laeven, Luc. 2012. The flight home effect: Evidence from the syndicated loan market during financial crises. *Journal of Financial Economics*, **104**(1), 23–43.

Giannola, Adriano, Lopes, Antonio, & Zazzaro, Alberto. 2013. La convergenza dello sviluppo finanziario tra le regioni italiane dal 1890 ad oggi. *Rivista di Politica Economica*, **103**(1), 145–197.

Goetz, Martin, Laeven, Luc, & Levine, Ross. 2016. Does the geographic expansion of banks reduce risk? *Journal of Financial Economics*, **120**(2), 346–362.

Greenwood-Nimmo, Matthew, Nguyen, Viet Hoang, & Shin, Yongcheol. 2015. *Measuring the connectedness of the global economy.* Working paper No. 7/15. The University of Melbourne.

Guarda, Paolo, Rouabah, Abdelaziz, & Theal, John. 2012. *An MVAR framework to capture extreme events in macro-prudential stress tests.* Working Paper Series 1464. European Central Bank.

Hsu, Nan-Jung, Hung, Hung-Lin, & Chang, Ya-Mei. 2008. Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis*, **52**(7), 3645–3657.

Illueca, Manuel, Norden, Lars, & Udell, Gregory F. 2013. *Do changes in the timeliness of loan loss recognition affect bank risk taking.* SSRN Working paper.

Imai, Masmai, & Takarabe, Seitaro. 2011. Bank integration and transmission of financial shocks: Evidence from Japan. *American Economic Journal: Macroeconomics*, **3**(1), 155–183.

Kaminsky, Graciela L, & Reinhart, Carmen M. 2000. On crises, contagion, and confusion. *Journal of international Economics*, **51**(1), 145–168.

Kock, Anders Bredahl, & Callot, Laurent. 2015. Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, **186**(2), 325–344.

Koop, Gary, Pesaran, M Hashem, & Potter, Simon M. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, **74**(1), 119–147.

Kwiatkowski, Denis, Phillips, Peter CB, Schmidt, Peter, & Shin, Yongcheol. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, **54**(1-3), 159–178.

Lütkepohl, Helmut. 2005. *New introduction to multiple time series analysis.* Springer-Verlag Berlin Heidelberg.

Meslier, Céline, Morgan, Donald P, Samolyk, Katherine, & Tarazi, Amine. 2016. The benefits and costs of geographic diversification in banking. *Journal of International Money and Finance*, **69**, 287–317.

Papi, Luca, Sarno, Emma, & Zazzaro, Alberto. 2015. *The geographical network of bank organizations: issues and evidence for Italy.* Working paper No. 105. Mo.Fi.R.

Park, Hail, & Shin, Yongcheol. 2017. Exploring international linkages using generalised connectedness measures: The case of Korea. *International Review of Economics & Finance*, **50**, 49–64.

Pesaran, H Hashem, & Shin, Yongcheol. 1998. Generalized impulse response analysis in linear multivariate models. *Economics letters*, **58**(1), 17–29.

Popov, Alexander, & Udell, Gregory F. 2012. Cross-border banking, credit access, and the financial crisis. *Journal of International Economics*, **87**(1), 147–161.

Presbitero, Andrea F, Udell, Gregory F, & Zazzaro, Alberto. 2014. The home bias and the credit crunch: A regional perspective. *Journal of Money, Credit and Banking*, **46**(s1), 53–85.

Pykhtin, Michael. 2004. Multi-Factor Adjustment. *Risk*, **17**(3), 85–90.

Schnabl, Philipp. 2012. The international transmission of bank liquidity shocks: Evidence from an emerging market. *The Journal of Finance*, **67**(3), 897–932.

Tibshirani, Robert. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tola, Vincenzo. 2010. *Measuring sectoral/geographic concentration risk.* Questioni di Economia e Finanza (Occasional papers) 72. Bank of Italy.

Van Rijckeghem, Caroline, & Weder, Beatrice. 2001. Sources of contagion: is it finance or trade? *Journal of international Economics*, **54**(2), 293–308.

Virolainen, Kimmo. 2004. *Macro stress testing with a macroeconomic credit risk model for Finland.* Discussion paper 18. Bank of Finland.

Zazzaro, Alberto. 2006. La scomparsa dei centri decisionali dal sistema bancario meridionale. *Rivista di Politica Economica*, **96**(2), 31–60.

Zou, Hui. 2006. The adaptive lasso and its oracle properties. *Journal of the American statistical association*, **101**(476), 1418–1429.

Zou, Hui, & Hastie, Trevor. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**(2), 301–320.

Zou, Hui, & Zhang, Hao Helen. 2009. On the adaptive elastic-net with a diverging number of parameters. *Annals of statistics*, **37**(4), 1733.

Chapter 3

Housing Market Shocks in Italy: a GVAR approach

3.1 Introduction

This chapter investigates the spatial and temporal diffusion of house prices and transaction volumes across 93 Italian provincial housing markets, over the period $2004 - 2016$. The transmission mechanism of house price spillovers across space and time is known in literature as "ripple effect". Meen (1999) gives four different explanations of the "ripple effect" in the UK housing markets – migration, equity transfer, spatial arbitrage and exogenous shocks. In particular, migration or equity transfer (e.g. longer-term residents of one area accumulate significant wealth in their home equity, cash out that equity by selling their home and moving to a lower cost region where a similar quality house costs much less) could lead to the ripple effect by increasing demand and thereby prices. Moreover, investors could spatially arbitrage their funds to acquire properties in lower priced regions, where higher anticipated returns exist on housing investment. In this case, financial capital moves, rather than households, between regions to link house prices. Finally, ripple effect pattern can be ascribed to heterogeneous responses of each region to exogenous macro conditions.
Empirical evidence of house prices spillovers across regions has been provided for UK (see Holly *et al.*, 2011; Gray, 2012; Tsai, 2014; Montagnoli & Nagayasu, 2015, among the others), for US (Brady, 2011, 2014), for China (Gong *et al.*, 2016) or for Denmark (Hviid, 2017). Most of these studies control for long run convergence in house prices by taking into account error correcting dynamics to long-run equilibrium relationship between house prices. The long-run analysis is particularly suitable to explore the role played by observed fundamentals (income and interest rates) in shaping the house prices

long-run dynamics. However, given the short time data span, my analysis does not control for long-run equilibrium and error correcting dynamics.
My first contribution to the existing studies on ripple effect using only house prices is based on an extension of the information set to transaction volumes in order to better capture the local housing market dynamics and the associated spillovers effect across space and time.[1]
Second, my analysis allows to assess heterogeneity in the spatial-temporal diffusion. While most of the studies on "ripple effect" focus on spillovers from a dominant unit, in this chapter I analyze how a specific shock to the house prices and transaction volumes arising from 10 Italian regional capitals spills over to other urban areas (their neighbours).
Finally, I contribute to the literature on the house price-volume correlation, which is based on the evaluation of the dynamic effects of observable housing market fundamentals on the price-volume co-movements (see Andrew & Meen, 2003; Clayton *et al.*, 2010, among the others). In particular, I analyze the spatio-temporal diffusion of house prices and volumes driven by unobserved fundamentals. The latent variable is interpreted as a negative housing demand shock identified through sign restrictions on house prices and transaction volumes modelled through a Global VAR, GVAR. The structural form impulse response analysis is informative on how local adverse shocks to fundamentals (which could be interpreted as a combination of negative income shock and a rise to interest rates) impact on house prices and volumes of the other areas (neighbours).
The GVAR model used for the empirical analysis, introduced by Pesaran *et al.* (2004), is a multi-country extension of the standard VAR model which allows to examine the temporal transmission of shocks within and between different geographical areas. The model allows to control for common factor effects, by using a spatial exogenous regressor, and, therefore, it allows to evaluate "genuine" spatial spillover effects across different housing markets. The structural housing demand shock is identified through theory-driven sign restrictions following the approach recently proposed in the study of Eickmeier & Ng (2015), which focuses on the transmission of US credit supply shocks across a panel of 33 countries over the period $1983 - 2009$.

In this chapter, I use semi-annual observations on real house prices and transaction volumes for 93 Italian provinces, over the period $2004 - 2016$. More specifically, I use a confidential and unique dataset provided by the Real Estate Market Observatory managed by the Italian Revenue Agency ("Agenzia delle Entrate - Osservatorio del Mercato Immobiliare") for the house prices. This rich dataset contains information at semi-annual frequency on maximum and minimum house prices (nominal, in euro)

[1] To my knowledge, the only study taking into account transaction volumes in estimating the "ripple effect" is the study of Tsai (2014).

categorized by types of real estate (housing, appurtenances, office, retail and industrial) and areas (i.e. central, suburbs, hinterlands), at municipal level, over the period from second semester 2002 to second semester 2016. As for the transaction volumes, I use quarterly observations for the number of normalized transaction (NNT), collected from the publicly available database of the Real Estate Market Observatory - Italian Revenue Agency ("Agenzia delle Entrate - Osservatorio del Mercato Immobiliare"), covering the 2004$Q1$ - 2016$Q4$ time span. To match the semi-annual data frequency of house prices, I aggregate the quarterly data on volumes, by taking the sum over two consecutive quarters.

My analysis provides some interesting findings. First, contrary to a large body of literature, this study does not find evidence of a "ripple effect" in house prices, with the notably exception of Rome. Second, I find evidence of a "ripple effect" in transaction volumes. In particular, the empirical results show that transaction volumes largely spill over across regional capitals and neighbours in response to the negative housing demand shock.

This chapter is structured as follows. Section 3.2 provides a literature review on "ripple effect" and the price-volume correlation. Section 3.3 describes the empirical methodology. Section 3.4 describes data and the empirical findings and Section 3.5 concludes.

3.2 Literature review

3.2.1 Spatio-temporal analysis of "ripple effect"

The "ripple effect" embodies two prominent feature of house price dynamics. The first is spatial dependence, e.g. cross-sectional correlation, relating each cross-section unit to its neighbours. The second one (which is fully accounted by recent empirical studies based on spatial autoregressive and spatial error component models) is the lagged transmission of price changes across neighbours, given that information takes time to travel, especially in a market for relatively illiquid assets.

Recently, Holly *et al.* (2011) compute impulse response analysis based on a Vector Autoregression (VAR) model (which includes a common spatial regressor as exogenous variable) fitted to house prices in London and 11 UK regions. The authors find evidence of dynamic house prices spillovers from London to neighbouring regions in the UK. Brady (2011), focusing on California counties, estimates spatial IRFs obtained from a single-equation spatial autoregressive panel model. The author, using the Jordà (2005) local projection method (involving direct forecasting techniques) to get the impulse response function, finds that a shock to an average county house prices in California has a positive (lasting two and half years) effect on the average house prices in a neighbouring region. Brady (2014) computes spatial IRFs for US states, obtained from the estimation of a single equation spatial autoregressive model for house prices, including state-specific covariates such as real income, interest rates and housing starts (and their lags). A central role in the single equation dynamic model used in both studies is played by the "spatial regressor" treated as exogenous variable. The spatial IRF analysis in Brady (2014) shows that a shock to housing prices at the state level has persistent effect (reaching the steady state within four years) on the panel of US states. The study of Gong *et al.* (2016) lends support to the house price temporal diffusion effect in a large emerging market such as China. The authors focus on monthly house price indexes of 10 cities' housing markets in the Pan-Pearl River Delta (Pan-PRD) area of China, covering the period from June 2005 to May 2015. The generalized impulse response functions (GIRF) obtained from traditional VAR (without spatial regressors) confirm a propagation of the house price shocks occurring to a given city approximately in accordance with the distance decay pattern found in the study of Holly *et al.* (2011). Hviid (2017), using a Global Vector Error Correction Model (VECM), augmented with a common spatial regressor, fitted to Danish house price data, finds strong evidence of a "ripple effect" in the short run of the model, but less so in the long run. This finding is interpreted as the "ripple effect" playing an important role as push factor in the short run, while house prices are mainly determined by regional fundamental factors in the long run.

All the aforementioned studies control for long run convergence in house prices (at least within clubs) by taking into account error correcting dynamics to long-run equilibrium relationship between house prices.

The study of Meen (1999) highlights the important role played by structural differences in regional housing markets (including different local economic conditions), beyond migration and spatial arbitrage. Therefore, the author suggests to focus on spatial coefficient heterogeneity when studying the dynamics of UK regional house prices. The study of Meen (1999) has inspired a number of researchers to analyze heterogeneity in the "ripple effect" (e.g. spatial heterogeneity). Van Dijk *et al.* (2011) detect the existence of two clusters of regions (mainly in terms of the average house prices growth rate) in the Netherlands: regions within the cluster have the same house price dynamics. Moreover, Gray (2012), using exploratory spatial data analysis and house price data from local authority districts in England and Wales, finds evidence that house price spillover north of the East Midlands appears much more rapid than what would be consistent with a "ripple effect". The empirical findings of the study suggest that there is some support for the analysis of the British housing market on a spatially segmented basis, even at a regional level. The study of Montagnoli & Nagayasu (2015) investigates the presence of house prices spillover among 12 UK regions over the period 1983Q1-2012Q3. The authors, using the approach proposed by Diebold & Yilmaz (2009) on VAR models fitted

to either 12 regional house price inflation rate or to the corresponding house price inflation volatility, find evidence of a "ripple effect" from London house prices to the other UK regions, whose magnitude declines as the spatial distance from London increases.[2]
Pijnenburg (2017), focusing on a balanced panel of 319 Metropolitan Statistical Areas (MSAs) of the US, observed over the period from 2004Q2 to 2009Q2, estimates a panel smooth transition regression model in order to capture the heterogeneity in spatial dependence across time and space as well as the heterogeneity in the effect of the fundamentals. The author finds evidence of heterogeneity spatial spillovers of house prices across space and time. In particular, heterogeneity in the effect of the fundamentals on house price dynamics is only found for population growth and building permits, but not for real per capita disposable income and the unemployment rate.

To my knowledge, the only study on the "ripple effect" taking into account transaction volumes is the one of Tsai (2014). The author, using monthly data on house prices and transaction volumes over the period 1995m2-2012m3, examines the presence of long-run convergence among 10 UK regional housing market. The use of panel-based unit root tests (developed by Im *et al.* (2003)) finds evidence of stationarity in the ratios of the regional to national house prices (as well as the one for transaction volumes). These findings are interpreted as evidence of convergence for both house prices and transaction volumes. Moreover, the analysis of Tsai (2014) shows that volumes converge to its equilibrium faster than the house prices.

3.2.2 House price-volume correlation

This chapter also seeks to contribute to the literature on the relationship between prices and transaction volumes and the underlying housing market fundamentals.
An early studies of Follain & Velz (1995) for the US and the one of Hort (2000) for Sweden suggest a negative relationship between house prices and sales volumes. In particular, Hort (2000) investigates to what extent a housing demand shock impacts on house prices and transaction volumes in the Swedish regional housing markets. Using data on house prices, transaction volumes and after-tax mortgage rate, the author employs VAR using monthly data (over the $1981 - 1993$ time data span) or quarterly data (over the $1982 - 1996$ sample period). The empirical findings (especially those based on monthly observations) reveal a strong negative reaction of sales, on impact, to a positive shock on nominal interest rate, while house prices start to decrease after $3 - 4$ months.
Other more recent empirical studies point at a positive correlation between the two covariates. Andrew & Meen (2003), using data for UK house prices and transaction

[2] In a first stage of the analysis, Montagnoli & Nagayasu (2015) test the UK regional house prices convergence, finding evidence of four convergence clubs.

volumes over the period $1969 - 1996$, estimate the adjustment mechanism of the two variables to their fundamentals. First, the authors construct a measure of long-run housing market disequilibrium (defined as the ratio of the desired owner-occupier housing stock to the actual stock), by using housing market fundamentals variables, such as income, housing stock, number of households and construction costs. In a second stage of the analysis, the authors estimate a conditional VAR model where house prices and turnover rate are regressed on deviations from equilibrium. Their findings show a positive correlation between house price and volume in the short-run period. Further, the authors find that volumes exhibit an adjustment faster than the house price in reaction to a shock to fundamentals. Empirical evidence of a positive price-volume correlation is also provided by the study of Clayton *et al.* (2010). Using data for 114 MSA of the US observed over the $1990 - 2002$ period, the authors estimate a panel VARX fitted to house prices and turnover rates. The exogenous variables are covariates related to labour market conditions and they are used as proxies of fundamentals. The authors show that the positive co-movement of the housing market aggregates is mainly driven by shocks to employment and household's income. Moreover, the authors find that transaction volumes react more than house prices to exogenous shocks. De Wit *et al.* (2013) focus on the Dutch economy (the sample period considered is $1985 - 2007$) and they use a Vector Error Correction model (VECM) fitted to proxies of house price (the real list price and the real transaction price), proxies of volume (the rate of entry and the rate of sale) and proxies of housing market fundamentals (unemployment and the real mortgage interest rate). The authors find evidence of an interest shock reducing both house prices and transaction volumes.

A number of studies provide some theoretical support to the evidence of a positive correlation between house prices and transaction volumes.[3] The positive price-volume correlation might depend on the presence of financial constraints. For example, the study of Stein (1995) develops a model where a positive shock to the housing market fundamentals increases prices as well as producing more incentive in demanding house, with an increase in the entry of new houses for sales, hence in the transaction volumes. Other studies have stated that the empirical evidence on the positive relationship between house prices and sales can be explained with the use of a search model where the idiosyncratic preferences of potential buyers are modelled on the basis of a mismatch costs between buyers and sellers (see Berkovec & Goodman, 1996, among the others). Finally, the positive correlation between house prices and transaction volumes might be caused by the market liquidity. In particular, Krainer (2001) considers a model of individual choice under uncertainty and frictions, where buyer and seller's decisions are jointly modelled. In equilibrium, both sellers and buyers maximize their expected values

[3] De Wit *et al.* (2013) provide an extended review of the main theoretical frameworks.

in a price-setting model. When the market price is high, "the opportunity cost of keeping an empty house" on the market increases, because the value of the house might decrease in the next period. In such a context, sellers slightly decrease their reservation price, matching the one that buyers are willing to pay, and the transaction volume increases.

While the aforementioned studies focus on observed fundamentals as drivers of price-volume co-movement, in this study I focus on unobserved fundamentals identified on housing demand shock.
Recently, a number of studies have focused on the role played by unobservable fundamentals, that is a housing demand shock in driving the international transmission of house price across countries. Vansteenkiste & Hiebert (2011) analyze the house price spillover mechanism across 7 Euro area countries, over the $1971 - 2009$ time span. The empirical model used is a Global VAR, GVAR, fitted to real house prices, real per capita income, and real long-term interest rate. Vansteenkiste & Hiebert (2011) find evidence of heterogeneity in the relatively small country house price responses to demand shocks. Cesa-Bianchi (2013) examines the international transmission of housing demand shocks using data on 33 Advanced Economies (AEs) and Emerging Market Economies (EMEs), for the period $1983 - 2009$. The author uses a GVAR model to evaluate to what extent a housing demand shock in US impact on a set of macroeconomic and financial variables, including GDP and house price. In a second stage of the analysis, the author estimates the GDP response to regional housing demand shocks (a synchronized increase in house prices in AEs). Although the main focus of the paper is on the response of the GDP across countries, Cesa-Bianchi (2013) finds that an increase in house prices also affects foreign housing markets.

3.3 Empirical methodology

3.3.1 The GVAR Model

The Global Vector Autoregression (GVAR) model was formerly introduced by Pesaran *et al.* (2004). This model is a multi-country extension of the VAR model and it allows to examine the interdependencies between cross-section units (say countries or provinces, for example) as well as assessing the spillover effects among the entire global system.
One advantage of modelling a GVAR model is that it addresses the problem of dimensionality, in particular when the number of endogenous variables for each cross-section unit becomes relatively large.

The main idea behind the GVAR model is the connection between country-specific variables (say domestic variables), y_{it}, and foreign country-specific variables, y^*_{it}.
The construction of the GVAR requires two steps.
The first step consists of linking the vector of country-specific variables, y_{it}, to the global economy variables, which comprise of the foreign country-specific variables and the deterministic components (see Pesaran *et al.*, 2004; Chudik & Pesaran, 2016), through the estimation of individual small-scale country-specific models.
Given N countries (cross-section units), indexed by $i = 1, \ldots, N$, each i-th small-scale country-specific model is shaped as a Vector Autoregression process for the $k_i \times 1$ vector of domestic variables, $y_{it} = (y_{i1}, \ldots, y_{iT})'$, augmented with the $k^*_i \times 1$ vector of foreign variables, $y^*_{it} = (y^*_{i1}, \ldots, y^*_{iT})'$, their lagged variables and the deterministic components.[4]
For the i-th country, the $VARX^*(p_i, q_i)$ model has the following reduced form representation:

$$y_{it} = a_{i0} + a_{i1}t + \sum_{\ell=1}^{p_i} \Phi_{i\ell} y_{i,t-\ell} + \Lambda_{i0} y^*_{it} + \sum_{\ell=1}^{q_i} \Lambda_{i\ell} y^*_{i,t-\ell} + u_{it} \tag{3.1}$$

for $i = 1, \ldots, N$ and for $t = 1, \ldots, T$, where $\Phi_{i\ell}$, for $\ell = 1, \ldots, p_i$, are the $k_i \times k_i$ coefficients matrices associated with the lagged endogenous variables, Λ_{i0} and $\Lambda_{i\ell}$, for $\ell = 1, \ldots, q_i$, are the $k_i \times k^*_i$ coefficients matrices associated with the foreign country-specific variables, a_{i0} is a $k_i \times 1$ vector of constant terms, a_{i1} is the $k_i \times 1$ vector of time trend coefficients and u_{it} is a $k_i \times 1$ vector of country-specific reduced form residuals, with zero mean and a nonsingular covariance matrix, that is $u_{it} \sim iid(0, \Sigma_{u_i})$.
The estimation of the $VARX^*$ model allows for conditioning the vector of the endogenous variables, y_{it}, to the foreign country-specific variables, y^*_{it}, which can be considered as a proxy of the global economy dimension.
Before discussing the construction of the country-specific foreign variables, let me define $z_{it} = (y_{it}, y^*_{it})'$ as the $(k_i + k^*_i) \times 1$ vector containing both domestic and foreign variables. Once fixing a maximum lag order, $r_i = max(p_i, q_i)$, the eq.(3.1) can be written as follows:

$$A_i z_{it} = a_{i0} + a_{i1}t + \sum_{\ell=1}^{r_i} B_{i\ell} z_{i,t-\ell} + u_{it} \tag{3.2}$$

where

$$A_i = (I_{k_i}, -\Lambda_{i0}) \quad \text{and} \quad B_{i\ell} = (\Phi_{i\ell}, \Lambda_{i\ell}) \tag{3.3}$$

[4]The $VARX^*$ specification allows for the inclusion of global (weakly) exogenous variables (see Pesaran *et al.*, 2004, for further details).

where A_i and $B_{i\ell}$, for $\ell = 1, \ldots, r_i$, are $k_i \times (k_i + k_i^*)$ matrices of coefficients constructed from the estimation of the model in eq.(3.1).[5]

The country-specific foreign variables, y_{it}^*, included in z_{it}, are computed as weighted cross-sectional averages of all the domestic variables by adopting a “link” matrix, W_i, different for each i-th country, as follows:

$$z_{it} = W_i y_t \tag{3.4}$$

where $y_t = (y_{1t}', y_{2t}', \ldots, y_{Nt}')'$ is the $K \times 1$ stacked vector, with $K = \sum_{i=1}^{N} k_i$, containing all the endogenous variables in the N-dimensional panel, and W_i is a $(k_i + k_i^*) \times K$ matrix of fixed weights, w_{ij}, which capture the relationship between the N observed countries (see Section 3.3.2, for further details).[6]

In line with Pesaran *et al.* (2004), the link matrices, W_i, can be written as follows:

$$W_i = \begin{pmatrix} 0 & \ldots & I_{k_i} & \ldots & 0 \\ w_{i1} I_{k_i^*} & \ldots & w_{ii} I_{k_i^*} & \ldots & w_{iN} I_{k_i^*} \end{pmatrix} \tag{3.5}$$

The link matrix allows to aggregate the small-scale country-specific models in a more compact representation. In fact, once estimating the coefficients matrices for each of the N country-specific models in eq.(3.1), the second step in the GVAR strategy consists of rewriting the small-scale models, by combining eqs.(3.2) and (3.4), as:

$$A_i W_i y_t = a_{i0} + a_{i1} t + \sum_{\ell=1}^{r_i} B_{i\ell} W_i y_{t-\ell} + u_{it} \tag{3.6}$$

The Global VAR(r) model is obtained by stacking the N-dimensional model equations into a single model:

$$G y_t = a_0 + a_1 t + \sum_{\ell=1}^{r} H_\ell y_{t-\ell} + u_t \tag{3.7}$$

with

[5]In a GVAR framework, the country-specific $VARX^*$ models can be rewritten using an error correction representation (see Chudik & Pesaran, 2016, among the others). The use of an error-correction form would require a test on the weak exogeneity of the foreign variables y_{it}^* with respect to the long-run parameters. Since I do not take into account error correcting dynamics, in the rest of the analysis I simply treat the foreign variables, y_{it}^*, as exogenous.

[6]Recently, a number of authors have introduced time-varying weights in modelling GVARs (see Cesa-Bianchi *et al.*, 2014, among the others).

$$G = \begin{pmatrix} A_1 W_1 \\ A_2 W_2 \\ \vdots \\ A_N W_N \end{pmatrix}, \quad H_\ell = \begin{pmatrix} B_{1\ell} W_1 \\ B_{2\ell} W_2 \\ \vdots \\ B_{N\ell} W_N \end{pmatrix}, \tag{3.8}$$

and

$$a_0 = \begin{pmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{N0} \end{pmatrix}, \quad a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{pmatrix}, \quad u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Nt} \end{pmatrix}$$

where G and H_ℓ, for $\ell = 1, \ldots, r$, are the $K \times K$ matrices containing the estimated parameters of the N small-scale $VARX^*$ models, a_0 and a_1 are the stacked $K \times 1$ vectors containing intercepts and time trends coefficients, respectively, and u_t is the $K \times 1$ stacked vector of country-specific residuals, which are normally distributed with zero mean and a $K \times K$ covariance matrix which has the following representation:

$$\Sigma_u = \begin{bmatrix} \Sigma_{u_1} & \Sigma_{u_1 u_2} & \ldots & \Sigma_{u_1 u_N} \\ \Sigma_{u_2 u_1} & \Sigma_{u_2} & \ldots & \Sigma_{u_2 u_N} \\ \vdots & \vdots & \ddots & \ldots \\ \Sigma_{u_N u_1} & \Sigma_{u_N u_2} & \ldots & \Sigma_{u_N} \end{bmatrix} \tag{3.9}$$

where Σ_{u_i} is the covariance matrix of the reduced form residuals of the i-th country-specific model, while $\Sigma_{u_i u_j}$ is the covariance matrix of the reduced form residuals of the i-th and j-th country-specific models. If G is invertible, a standard VAR process for the stacked vector, y_t, can be easily obtained by pre-multiplying the elements in eq.(3.7) by G^{-1}:

$$y_t = b_0 + b_1 t + \sum_{\ell=1}^{r} F_\ell y_{t-\ell} + \epsilon_t \tag{3.10}$$

where $b_0 = G^{-1} a_0$, $b_1 = G^{-1} a_1$, $F_\ell = G^{-1} H_\ell$, for $\ell = 1, \ldots, r$, and $\epsilon_t = G^{-1} u_t$.
The eq.(3.10) can be solved recursively to conduct dynamic analysis, such as Impulse response, Forecast error variance decomposition and Historical decomposition analysis (Pesaran *et al.*, 2004).

The GVAR model is particularly useful in studying the propagation mechanism of shocks

among a relevant number of countries (provinces in my analysis). In fact, the use of a GVAR model allows to solve the problem of an over-parametrization which may occur in the estimation of standard VARs. However, the computational and data analysis challenges might raise as the number of cross-section units increases. As suggested by Pesaran *et al.* (2004), one possible solution is to aggregate cross-section units into blocks of regions, whose regional-specific $VARX^*$ models are shaped by using weighted averages of the variables for all the units that belong to a specific region.
Nonetheless, I prefer treating each province as a separate cross-section unit and construct provincial-specific $VARX^*$ models. Furthermore, the small-scale $VARX^*$ models are then aggregated into different GVARs, one for each of the 5 Italian groups of regions (say macro-regions): Northwest, Northeast, Central, South and Insular Italy (see Table 3.1).

3.3.2 Weighting strategy

The estimation of the province-specific $VARX^*(1,1)$ models involves the construction of the N weights matrices, W_i, with $i = 1, \dots, N$, in order to get spatial regressors by averaging out the foreign variables, $y^*_{it} = \sum_{j=1}^{N} w_{ij} y_{jt}$ (see eq.(3.5)).
Generally, the weighting strategy is modelled by using shares of cross-country trade flows (Pesaran *et al.*, 2004; Dees *et al.*, 2007a; Cesa-Bianchi, 2013, among the others), cross-country bank lending exposures (Galesi & Sgherri, 2009, among the others), a combination of weights based on both trade and financial flows (Eickmeier & Ng, 2015) or spatial-based weights (Vansteenkiste, 2007; Vansteenkiste & Hiebert, 2011).
Since my aim is to highlight a spatial dimension of the propagation mechanism of housing market shocks in the Italian provinces, I use time-fixed geographic weights.[7]
In order to construct the $(k_i + k^*_i) \times K$ province-specific link matrix, W_i, my methodology relies on constructing spatial weights based on contiguity between province i and province j (see Holly *et al.*, 2011).[8] Given N geographical units, the spatial matrix, labelled as S, is a $N \times N$ binary matrix with generic entries $w_{ij} \geq 0$, where $w_{ij} = 1$ if provinces i and j share a border and zero otherwise:

[7]Most of the data used in Spatial econometrics are on irregular areas, such as regions or provinces. Generally, information on irregular areas take the form of *shape files*, which include, for example, the spatial coordinates and the attributes associated to each spatial unit. In my analysis, I use the *shape file* downloaded from the Italian National Institute of Statistic (ISTAT), containing spatial information for the Italian provinces. The construction of the spatial weights matrix is implemented by using the **spdep** and **maptools** packages in R.

[8]The spatial weights can be also constructed on the basis of geographic distance (see Vansteenkiste, 2007; Vansteenkiste & Hiebert, 2011, among the others) or socio-economic distance (see Conley & Topa, 2002, for example).

Table 3.1: Italian provinces grouped at NUTS1 (macro-regional) and NUTS2 (regional) levels.

Macro-regions	Regions	Provinces
North-West	Aosta Valley	Aosta
	Liguria	Genoa, Imperia, La Spezia and Savona
	Lombardy	Bergamo, Brescia, Como, Cremona, Lecco, Lodi, Mantova, Milano, Pavia, Sondrio and Varese
	Piedmont	Alessandria, Asti, Biella, Cuneo, Novara, Torino Verbania and Vercelli
North-East	Emilia-Romagna	Bologna, Ferrara, Forlí-Cesena, Modena, Parma, Piacenza, Ravenna, Reggio Emilia and Rimini
	Friuli Venezia-Giulia	Pordenone and Udine
	Veneto	Belluno, Padova, Rovigo, Treviso, Venezia, Verona and Vicenza
Centre	Lazio	Frosinone, Latina, Rieti, Roma and Viterbo
	Marche	Ascoli Piceno, Ancona and Pesaro (and Urbino)
	Tuscany	Arezzo, Firenze, Grosseto, Livorno, Lucca, Massa (and Carrara), Pisa, Pistoia, Prato and Siena
	Umbria	Perugia and Terni
South	Abruzzo	Chieti, Pescara and Teramo
	Apulia	Bari, Brindisi, Foggia, Lecce and Taranto
	Basilicata	Matera and Potenza
	Calabria	Cosenza, Catanzaro, Crotone, Reggio Calabria and Vibo Valentia
	Campania	Avellino, Benevento, Caserta, Napoli and Salerno
	Molise	Campobasso and Isernia
Islands (or Insular)	Sicily	Agrigento, Caltanissetta, Catania, Enna, Messina, Palermo, Ragusa, Siracusa and Trapani

Note. Since the presence of missing values, I exclude provinces in Trentino Alto-Adige (a region of the North-East of Italy) and Sardinia (a region of Insular Italy) from the analysis (see Section 3.4.1).

$$S = \begin{pmatrix} 0 & w_{12} & \dots & w_{1N} \\ w_{21} & 0 & \dots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & \dots & \dots & 0 \end{pmatrix} \tag{3.11}$$

with $w_{ij} = w_{ji}$. Note that the main diagonal elements in S are zero, $w_{ii} = 0$, by construction. Furthermore, I standardize S by row sum ($\bar{S}$), with generic entries $\bar{w}_{ij} = 1/n_i$, where n_i is the number of neighbours of the i-th province (see also Holly *et al.*, 2011). These spatial weights are then rearranged into the link matrices, W_i (see eq.(3.5)).

3.3.3 Estimation procedure

For each of the 5 Italian macro-regions, I construct a bivariate GVAR model (with no time trend) for real house price (ΔHP) and sales (ΔNTN) changes, where the corresponding province-specific $VARX^*$ models present a lag of order one for both domestic and foreign variables, $VARX^*(1,1)$:

$$y_{it} = a_{i0} + \Phi_{i1} y_{i,t-1} + \Lambda_{i0} y^*_{it} + \Lambda_{i1} y^*_{i,t-1} + u_{it} \tag{3.12}$$

for $i = 1, \dots, N$ and for $t = 1, \dots, T$. The model in eq.(3.12) can be written as:

$$A_i W_i y_t = a_{i0} + B_{i1} W_i y_{t-1} + u_{it} \tag{3.13}$$

where $A_i = (I_{k_i}, -\lambda_{i0})$, $B_{i1} = (\Phi_{i1}, \lambda_{i1})$ and $W_i y_t = z_{it}$ (see eqs.(3.1)-(3.6)). Further, according to Table 3.1, the province-specific models are rearranged into the corresponding GVAR models:

$$G y_t = a_0 + H_1 y_{t-1} + u_t \tag{3.14}$$

and

$$y_t = b_0 + F_1 y_{t-1} + \epsilon_t \tag{3.15}$$

where $b_0 = G^{-1} a_0$, $F_1 = G^{-1} H_1$ and $\epsilon_t = G^{-1} u_t$ (see eqs.(3.7)-(3.10)). Therefore, the $K \times K$ coefficient matrix associated to the stacked vector of endogenous variables at time $t-1$ can be expressed as a function of the estimated province-specific $VARX^*(1,1)$ models:

$$F_1 = G^{-1} H_1 = \begin{pmatrix} (I_{k_1}, -\lambda_{10}) W_1 \\ (I_{k_2}, -\lambda_{20}) W_2 \\ \vdots \\ (I_{k_N}, -\lambda_{N0}) W_N \end{pmatrix}^{-1} \times \begin{pmatrix} (\Phi_{11}, \lambda_{11}) W_1 \\ (\Phi_{21}, \lambda_{21}) W_2 \\ \vdots \\ (\Phi_{N1}, \lambda_{N1}) W_N \end{pmatrix} \tag{3.16}$$

3.3.4 Structural Identification

Generally, the identification of shocks in GVARs is based on the Generalized impulse response functions (GIRF) framework originally proposed by Koop *et al.* (1996). Although this approach is not sensitive to the ordering of the variables, it admits correlated errors, hence the economic interpretation of the resulting shocks might be difficult (see Pesaran *et al.*, 2004). More recently, a number of studies have extended structural identification schemes to GVARs, to identify a single shock or a subset of shocks through either a Cholesky factorization (see Dees *et al.*, 2007a; Cesa-Bianchi, 2013, among the others) or through sign restrictions (Chudik & Fidora, 2011).
Recently, Eickmeier & Ng (2015) introduce a novel approach where the structural shocks are identified by imposing sign restrictions on the impulse responses obtained from a GVAR model.

3.3.4.1 Focus on identification through sign restrictions in VARs

In the last few years, the use of identification through sign restrictions in standard VARs has become a popular tool as an alternative to the traditional identification schemes, such as short-run restrictions, long-run restrictions introduced by Blanchard & Quah (1989) and the identification through heteroskedasticity approach (see Rigobon, 2003; Lanne & Lütkepohl, 2008, among the others). This approach is essentially based on the imposition of signs on the responses of the variables to one or more structural shocks, on the basis of economic theories (see Faust, 1998; Uhlig, 2005, among the others).
To explain how the identification through sign restrictions is implemented, consider a k_i-dimensional structural Vector Autoregression model for country i of order p:[9]

$$A_0^{(i)} y_t^{(i)} = A_1^{(i)} y_{t-1}^{(i)} + \ldots + A_p^{(i)} y_{t-p}^{(i)} + \varepsilon_t^{(i)} \tag{3.17}$$

where $A_0^{(i)}$ is a $k_i \times k_i$ matrix which contains the contemporaneous relationship among the variables, $A_\ell^{(i)}$, for $\ell = 1, \ldots, p$, are the coefficients matrices associated to the lagged variables and $\varepsilon_t^{(i)}$ are the zero-mean structural disturbances, with a covariance matrix, $\Sigma_\varepsilon^{(i)} = E(\varepsilon_t^{(i)} \varepsilon_t^{(i)'}) = I_{k_i}$, which is assumed to be an identity matrix. The corresponding reduced form representation can be easily obtained by pre-multiplying $(A_0^{(i)})^{-1}$ in eq.(3.17):

[9] In this subsection, the superscript i is used in the notation to indicate that I refer to a single country model.

$$\begin{aligned}(A_0^{(i)})^{-1}A_0^{(i)}y_t^{(i)} &= (A_0^{(i)})^{-1}A_1^{(i)}y_{t-1}^{(i)} + \ldots + (A_0^{(i)})^{-1}A_p^{(i)}y_{t-p}^{(i)} + (A_0^{(i)})^{-1}\varepsilon_t^{(i)} \\ y_t^{(i)} &= \Phi_1^{(i)}y_{t-1}^{(i)} + \ldots + \Phi_p^{(i)}y_{t-p}^{(i)} + u_t^{(i)}\end{aligned} \tag{3.18}$$

with $(A_0^{(i)})^{-1}A_\ell^{(i)} = \Phi_\ell^{(i)}$, for $\ell = 1, \ldots, p$, and $(A_0^{(i)})^{-1}\varepsilon_t^{(i)} = u_t^{(i)}$. Therefore, $\Sigma_u^{(i)} = E(u_t^{(i)}u_t^{(i)\prime}) = E\big[(A_0^{(i)})^{-1}\varepsilon_t^{(i)}\varepsilon_t^{(i)\prime}(A_0^{(i)})^{-1\prime}\big] = (A_0^{(i)})^{-1}\Sigma_\varepsilon^{(i)}(A_0^{(i)})^{-1\prime}$ denotes the relationship between the structural and reduced form covariance matrices of residuals. Since $\Sigma_u^{(i)} = P^{(i)}P^{(i)\prime}$, where $P^{(i)}$ is the $k_i \times k_i$ lower-triangular Cholesky decomposition of $\Sigma_u^{(i)}$, and $\Sigma_\varepsilon^{(i)} = I_{k_i}$, it follows that $(A_0^{(i)})^{-1} = P^{(i)}$. One of the most popular scheme employed to orthogonalize the reduced form residuals consists of imposing zero restrictions directly on $(A_0^{(i)})^{-1}$, by using the Cholesky decomposition of $\Sigma_u^{(i)}$, $u_t^{(i)} = P^{(i)}\varepsilon_t^{(i)}$ (see Kilian & Lütkepohl, 2017, for further details). However, the economic interpretation of the resulting structural disturbances might be difficult to achieve.

Another possible way is to randomly generate a set of candidate solutions, $\varepsilon_t^{(i)^*}$, for the structural disturbances, $\varepsilon_t^{(i)}$, where each solution has mutually uncorrelated shocks with unity variance:

$$\varepsilon_t^{(i)^*} = Q'\varepsilon_t^{(i)} \tag{3.19}$$

where Q is a $k_i \times k_i$ orthonormal matrix, such that $Q'Q = QQ' = I_{k_i}$ and $u_t^{(i)} = P^{(i)}QQ'\varepsilon_t^{(i)} = P^{(i)}Q\varepsilon_t^{(i)^*}$. Once generating the set of candidate solutions, $\varepsilon_t^{(i)^*}$, the identification of the structural shocks relies on selecting the corresponding structural impact multiplier matrix, $P^{(i)}Q$, which respects the theory-driven sign restrictions imposed on $(A_0^{(i)})^{-1}$.

The basic idea of the identification through sign restrictions relies on generating candidates orthogonal matrices, Q, from a set of all orthogonal matrices, $\mathcal{O}$.

Kilian & Lütkepohl (2017) define the set of $k_i \times k_i$ orthogonal matrices, $\mathcal{O}(K)$, as:

$$\mathcal{O}(k_i) \equiv \{Q | Q'Q = I_{k_i}\} \tag{3.20}$$

In literature, there are two common approaches to generate orthonormal matrices including: the Givens rotation matrices method and the Householder transformation approach. Let me consider a VAR model with $k_i = 2$ endogenous variables, $y_t^{(i)} = (y_{1t}^{(i)}, y_{2t}^{(i)})'$, and two structural shocks, $\varepsilon_t^{(i)} = (\varepsilon_{1t}^{(i)}, \varepsilon_{2t}^{(i)})'$. The relationship between the reduced form residuals and the structural disturbances is $u_t^{(i)} = (A_0^{(i)})^{-1}\varepsilon_t^{(i)}$:

$$\begin{pmatrix} u_{1t}^{(i)} \\ u_{2t}^{(i)} \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{pmatrix} \varepsilon_{1t}^{(i)} \\ \varepsilon_{2t}^{(i)} \end{pmatrix} \tag{3.21}$$

The Givens rotation matrices method consists of generating $k_i \times k_i$ matrices which have the following form:

$$Q(\phi) = \begin{bmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{bmatrix} \tag{3.22}$$

with $\phi \in [0, 2\pi]$. Therefore, the set of $k_i \times k_i$ orthogonal matrices can be defined as:

$$\mathcal{O}(k_i) = \{Q(\phi) | \phi \in [0, 2\pi]\} \tag{3.23}$$

To generate the orthogonal matrix $Q(\phi)$, one can define a finite-dimensional grid over the possible values of ϕ or randomly draw ϕ from a uniform distribution on $[0, 2\pi]$. Once obtaining a draw for $Q(\phi)$ and the corresponding structural impact multiplier matrix, $P^{(i)}Q(\phi)$, one can keep the rotation whose the associated impulse response respects the sign restrictions (see Kilian & Lütkepohl, 2017).

An alternative approach, known in literature as Householder transformation, is the one based on the QR decomposition proposed by Rubio-Ramirez *et al.* (2010). Given a $k_i \times k_i$ real square matrix, X, the associated QR decomposition is given by $X = QR$, where Q is a $k_i \times k_i$ orthogonal matrix and R is a $k_i \times k_i$ upper triangular matrix. If X is invertible, restricting the diagonal elements of R to be positive ensures that the QR decomposition of X is unique (see Stewart, 1980, for further details).

In a first step, the algorithm proposed by Rubio-Ramirez *et al.* (2010) consists of drawing a $k_i \times k_i$ matrix, X, with each element having an independent standard gaussian distribution.[10] In a second step, the authors suggest to compute its QR decomposition, $X = QR$, and normalize the main diagonal elements of R to be positive. According to the authors, these two steps correspond to draw Q from a uniform distribution over the set of orthogonal matrices $\mathcal{O}(k_i)$. Therefore, one can generate a large number of random draws, compute the corresponding structural impact multiplier matrices, $P^{(i)}Q$, and keep only those which agree with the imposed sign restrictions.

[10]The draw of each element entering X from an independent normal standard distribution ensures that X is invertible (see Kilian & Lütkepohl, 2017).

3.3.4.2 Identification through sign restrictions in GVAR: housing market shock

To identify the housing market shock, I follow the suggestions of Eickmeier & Ng (2015) relying on sign restrictions on the impulse responses obtained from a GVAR model. The generation of candidate structural impulse response relies on the Householder transformation approach proposed by Rubio-Ramirez *et al.* (2010), discussed in Section 3.3.4.1. Before commenting on the strategy used to identify the structural shocks, it is important to observe that the spillover analysis in GVARs, which captures the transmission mechanism of shocks across different cross-sectional units, relies on designating one unit (province in my analysis) as a "dominant unit".[11]

The structural identification strategy I adopt in the analysis requires the orthogonalization of the shocks originating from the dominant unit (or units), while it admits for shocks which are correlated with those originating from the remaining units.

Following Eickmeier & Ng (2015), the first step consists of computing the Cholesky decomposition of the N residuals covariance matrices, $\Sigma_{u_i} = E(u_{it}u_{it}')$, for $i = 1, \ldots, N$, obtained from the estimation of the individual reduced form province-specific $VARX^*$ models and, then, I combine the resulting Cholesky decomposition matrices, P_i into a $K \times K$ block diagonal matrix, P:

$$P = \begin{pmatrix} P_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & P_i & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & P_N \end{pmatrix} \tag{3.24}$$

The P matrix in eq.(3.24) is then used for the purpose of orthogonalizing the residuals of GVAR, u_t, defined in eq.(3.14), as $v_t = (v_{1t} \cdots v_{it} \cdots v_{Nt})' = P^{-1}G\epsilon_t$, where v_t has dimension $K \times 1$. Note that the relationship between u_t and ϵ_t is defined as $G^{-1}u_t = \epsilon_t$ (see 3.15). Therefore, the h-step ahead impulse responses matrices (which have dimension $K \times K$) associated with the orthogonalized residuals, v_t, are given by $\Psi^h = F_1^h G^{-1} P$. The (i, j)-th element denotes the h-step ahead response of the i-th endogenous variable to a shock occurring in the j-th endogenous variable.

Let me define m dominant (main) units, with $m \in N$. Following Eickmeier & Ng (2015), for the m-th dominant unit, I randomly draw $k_m \times k_m$ independent standard gaussian matrices, $\tilde{X}_m$, where k_m denotes the number of endogenous variables for the

[11] Generally, the dominant cross-section unit is labelled as unit 0 (see Pesaran *et al.*, 2004; Chudik & Pesaran, 2016, for example). However, the GVAR model I consider for each macro-region includes more than one dominant unit in the analysis.

Table 3.2: GVAR models, acceptance rate and matrix dimension.

GVAR	Regional capital	Acceptance rate	Matrix dimension
Northwest	Torino	1036/1000	$k = 2$, $N = 24$, $K = 48$
	Genova	1086/1000	
	Milano	1317/1000	
Northeast	Venezia	1262/1000	$k = 2$, $N = 18$, $K = 36$
	Bologna	1162/1000	
Centre	Firenze	1174/1000	$k = 2$, $N = 20$, $K = 40$
	Roma	1313/1000	
South	Napoli	1116/1000	$k = 2$, $N = 22$, $K = 44$
	Bari	1012/1000	
Islands	Palermo	1396/1000	$k = 2$, $N = 9$, $K = 18$

Note. The acceptance rate indicates the number of rotation matrices drawn, $\tilde{Q}_m$, necessary to obtain the 1000 valid point estimates of impulse responses. The Table also provides some information helpful to understand the dimension of matrices described in Section 3.3.4.2. k is the number of endogenous variables for each of the 93 Italian provinces, N is the number of provinces for each macro-regional GVAR (see also Table 3.1) and $K = \sum_{i=1}^{N} k_i$.

m-th unit. Since in my analysis the number of endogenous variables is equal to two, for each province, I let k_m equal to two for the rest of the section.
Further, I compute the QR decomposition of $\tilde{X}_m$, that is $\tilde{X}_m = \tilde{Q}_m \tilde{R}_m$ (see Rubio-Ramirez *et al.*, 2010).[12] For each replication, I multiply the $2 \times K$ orthogonalized residuals of the dominant unit, v_{mt}, by the 2×2 orthogonal matrix, $\tilde{Q}_m$, to obtain the structural shock for the m-th dominant unit, $\eta_{mt} = (\tilde{Q}_m v_{mt})'$.
Since I impose sign restrictions on the impulse response only on impact, I remove the superscript h from the notation, for the rest of this subsection.
The corresponding impulse responses ($h = 0$) are computed as $\Theta_m = (\Psi_m \tilde{Q}_m')'$, where Ψ_m is the 2×2 block matrix, for the selected m-th unit, in the impulse response matrix, Ψ (which has dimension $K \times K$). I discard the rotation matrices whose multiplications by the impulse responses, Ψ_m, do not satisfy the sign restrictions. In particular, I check the sign restrictions by focusing on the 2×2 matrix, Θ_m, for the m-th "dominant" unit. I repeat the algorithm until I save 1000 valid rotation matrices, $\tilde{Q}_m$.
For each of the 5 macro-regional GVAR models, I select one or more "dominant" units which correspond to the main regional capitals (or provinces) under investigation. Therefore, I apply the above described algorithm for the selected "dominant" unit. Following Eickmeier & Ng (2015), I also report the acceptance rates of the rotation matrices which

[12] To ensure that the QR decomposition of the independent standard gaussian matrix is unique, the diagonal of the upper triangular matrix, $\tilde{R}_m$, is normalized to be positive (see also Arias *et al.*, 2014).

satisfy the sign restrictions (see Table 3.2).[13] To better explain how the above described algorithm works, in Table 3.2 I also report some information on the dimension of matrices for each macro-regional GVAR model.

My focus is, first, on the identification of a negative innovation to housing demand in a specific regional capital (which is related to a combination of negative shock to income and a positive shock to interest rates), and also on the propagation of this shock to house prices and transaction volumes across neighbouring Italian provinces. For this purpose, I impose, on impact, a negative response both for house prices and transaction volumes (see Table 3.3).

The identification of only one structural shock in a system with two endogenous variables (house prices and sales) implies the estimation of a "partially identified" VAR (or $VARX^{*}$) model (see Kilian & Lütkepohl, 2017). To overcome the partial identification issue in a GVAR framework, I follow the suggestion reported by the study of Eickmeier & Ng (2015) which concentrates only on the identification of US credit supply shock.

Let me consider a generic m dominant unit. For each draw, I focus on the 2×2 block matrix $\Theta_m = (\Psi_m \tilde{Q}_m')'$ at zero horizon and I check if the response of the variables agrees with the sign restrictions in Table 3.3 as:

$$\Theta_m = \begin{bmatrix} \leq & \leq \\ n.a & n.a \end{bmatrix} \tag{3.25}$$

It is important to observe that the structural shock is reported in a row by construction. However, it is possible that a generic draw leads to the following situation:

$$\Theta_m = \begin{bmatrix} \leq & \leq \\ \leq & \leq \end{bmatrix} \tag{3.26}$$

where both of the two shocks are orthogonal, but their economic interpretation become difficult. Since I focus on identifying one structural shock (e.g. a negative housing demand shock), Eickmeier & Ng (2015) suggest to check the sign of the responses also in the second row and keep the draw if the signs in the second row are complement of the ones in the first row (see also Kilian & Lütkepohl, 2017):

$$\Theta_m = \begin{bmatrix} \leq & \leq \\ \leq & \geq \end{bmatrix} or \begin{bmatrix} \leq & \leq \\ \geq & \leq \end{bmatrix} \tag{3.27}$$

[13]The analysis of this chapter is conducted in R. My codes are, to a large extent, an adaptation of the Eickmeier & Ng (2015)'s MATLAB codes and the Galesi & Smith (2014)'s GVAR toolbox.

Table 3.3: Sign restrictions on impact

	Volumes	Prices
Housing market shock	-	-

Note: The sign restrictions refer to a negative shock. The restrictions are imposed as $\leq$.

Hence, I discard the draw in which the responses of the variables to the structural shocks report the same signs, as in eq.(3.26), otherwise I keep the draw. As mentioned before, I repeat the algorithm until I save 1000 valid rotation matrices, for each m-th dominant unit.

3.3.4.3 Median Target (MT) approach

In the last few years, the use of theory-driven sign restrictions has become a valid alternative tool for the identification of the structural shocks in VAR models. However, there are drawbacks associated with the use of sign restrictions. As argued by Fry & Pagan (2007), there is no guarantee that the impulse responses, which satisfy the imposed sign restrictions, come from the same model. Therefore, reporting the uncertainty of the identified impulse responses through the use of their quantiles might lead to wrong conclusions.

In line with Eickmeier & Ng (2015), once obtaining the set of impulse responses satisfying the sign restrictions, I apply the Median Target (MT) approach originally proposed by Fry & Pagan (2007), which is based on selecting the impulse responses which are the closest to the median values of those generated by all the admissible models.

According to the MT approach proposed by Fry & Pagan (2007), for each saved draw, that is $\tilde{Q}_m^{(r)}$, with $r = 1, \ldots, 1000$, I first standardize the associated identified h-step ahead impulse responses of the "dominant" unit by subtracting their median and dividing by their standard deviation. Further, since I only focus on the response on impact, I vectorize the 2×2 block matrix of impulse responses at $h = 0$, that is $\Theta_m^{(r)\prime} = (\Psi_m^{(r)} \tilde{Q}_m^{(r)\prime})$, in a 4×1 vector, $\theta_m^{(r)}$.

Finally, once selecting the r-th draw that minimizes $\theta_m^{(r)\prime} \theta_m^{(r)}$, say $\tilde{Q}_m^*$, I select from the $(K \times K)$ h-step ahead impulse responses matrix, Ψ^h, a $2 \times K$ matrix, ψ_m^h, for the m-th "dominant" unit and I multiply this matrix by $\tilde{Q}_m^*$, as $(\psi_m^{h\prime} \tilde{Q}_m^*)'$, for $h = 0, \ldots, H$, to produce the new set of impulse responses, which contains the responses of the K endogenous variables in the system to a shock occurring in the m-th "dominant unit".

Following Eickmeier & Ng (2015), $\tilde{Q}_m^*$ is also used to produce bootstrap for the GVAR model.

3.3.4.4 Bootstrapping the GVAR model

In particular, I use the sieve bootstrap procedure originally proposed by Bühlmann (1997) for Autoregressive (AR) processes and, more recently, employed by Dees *et al.* (2007a), Dees *et al.* (2007b) and Eickmeier & Ng (2015).
Following the approach reported in the study of Dees *et al.* (2007b), given the K-dimensional vector of residuals, $\hat{\epsilon}_t = (\hat{\epsilon}_{1t}, \hat{\epsilon}_{2t}, \ldots, \hat{\epsilon}_{Nt})'$, with $K = \sum_{i=1}^{N} k_i$, obtained from eq.(3.15), I randomly draw B series with replacement from the residuals, that is $\epsilon_t^{(b)} = (\epsilon_{1t}^{(b)}, \epsilon_{2t}^{(b)}, \ldots, \epsilon_{Nt}^{(b)})'$.[14]
To obtain the bootstrapped residuals $\epsilon_t^{(b)}$, I first pre-whiten the residuals $\hat{\epsilon}_t$ as $\hat{\eta}_t = \hat{A}^{-1}\hat{\epsilon}_t$, where $\hat{A}^{-1}$ is the generalized inverse obtained through a spectral decomposition of $\hat{\Sigma}_\epsilon$. In fact, the covariance matrix of the residuals $\hat{\epsilon}_t$ can be decomposed through a spectral decomposition, that is $\hat{\Sigma}_\epsilon = \hat{V}\hat{\lambda}\hat{V}'$, where $\hat{V}$ is an orthogonal matrix containing the eigenvectors, while $\hat{\lambda}$ is a diagonal matrix reporting the eigenvalues. The generalized matrix, $\hat{A}$, is then computed as $\hat{A} = \hat{V}\hat{\lambda}^{1/2}$.
The resampling with replacement is conducted on the pre-whiten residuals, $\hat{\eta}_t$.[15] For each b-th replication, with $b = 1, \ldots, B$, I compute the bootstrapped residuals of the GVAR model as $\epsilon_t^{(b)} = \hat{A}\hat{\eta}_t^{(b)}$ and I use them, together with the point estimates retrieved from eq.(3.15), to generate new artificial series, $y_t^{(b)}$:

$$y_t^{(b)} = \hat{b}_0 + \hat{F}_1 y_{t-1}^{(b)} + \epsilon_t^{(b)} \tag{3.28}$$

where $y_0^{(b)} = y_0$ are the actual initial observations. For each replication b, the artificial series are then used to retrieve new provincial-specific $VARX^*(1,1)$ estimates from:

$$y_{it}^{(b)} = \hat{a}_{i0}^{(b)} + \hat{\Phi}_{i1}^{(b)} y_{it-1}^{(b)} + \hat{\lambda}_{i0}^{(b)} y_{it}^{*(b)} + \hat{\lambda}_{i1}^{(b)} y_{it-1}^{*(b)} + \hat{u}_{it}^{(b)} \tag{3.29}$$

with $i = 1, \ldots, N$. From the estimation of the new provincial-specific $VARX^*(1,1)$ models in eq.(3.29), I construct the corresponding GVAR model and I compute the bootstrapped impulse responses.
In line with Eickmeier & Ng (2015), these impulse responses, identified once again through sign restrictions, are computed following the algorithm described above (see Section 3.3.4.2). However, differently from the point estimate, to check whether the

[14] I use the notation $\hat{\epsilon}_t$ to distinguish them from the bootstrapped residuals, $\epsilon_t^{(b)}$. However, it is important to note that $\hat{\epsilon}_t$ in eq.(3.15) are not directly estimated, since the estimation is conducted for the provincial-specific $VARX^*(1,1)$ models (see eq.(3.12)). In explaining the bootstrap procedure, I follow the same notation reported in Dees *et al.* (2007b) and I use the superscript "$(\hat{\cdot})$" to distinguish the quantities obtained through point estimation from the ones obtained by bootstrapping the GVAR.

[15] To reduce the complexity of the algorithm, Dees *et al.* (2007b) suggest to resample on a stacked version of the pre-whiten residuals $\hat{\eta}_t$.

impulse response of the m-th "dominant unit" respects the signs, I use the selected rotation matrix, $\tilde{Q}_m^{*\prime}$, to compute $\Theta_m^{(b)} = (\Psi_m^{(b)} \tilde{Q}_m^{*\prime})'$.
Finally, the $100(1-\alpha)\%$ confidence interval is constructed as $\alpha/2$ and $(1-\alpha)/2$ quantiles of the whole set of impulse responses to the identified structural shock, for each i-th province and h-th step ahead, on the basis of 200 bootstrap replications.

It is important to observe that in a GVAR model correlation between residuals arises within-country (e.g between the innovations associated with variables of a province-specific model) and across-countries (e.g. between the innovations to the same endogenous variable corresponding to different units, provinces). The identification through sign restrictions allows to addresses the issue of within-country residuals correlation. The issue of across-countries residuals correlation is addressed by conditioning the domestic endogenous variables, y_{it}, on the "foreign" variables, y_{it}^*. In order to check the cross-country correlation, I compute the average pairwise cross-country correlations among the endogenous variables and the individual $VARX^*(1,1)$ residuals (see Cesa-Bianchi, 2013; Eickmeier & Ng, 2015). Similar to the empirical findings of Cesa-Bianchi (2013) and of Eickmeier & Ng (2015), I obtain that the largest pairwise cross-country correlation between residuals (in absolute value) is 0.24, while the corresponding mean is 0.04 (see Table 3.4).

3.4 Empirical analysis

3.4.1 Data

I use semi-annual observations on real house prices and transaction volumes for 93 Italian provinces, over the sample period $2004-2016$.
More specifically, I use a confidential and unique dataset provided by the Real Estate Market Observatory managed by the Italian Revenue Agency ("Agenzia delle Entrate - Osservatorio del Mercato Immobiliare") for house prices. This rich dataset contains information at semi-annual frequency on maximum and minimum values of house prices (nominal, in euro) categorized by types of real estate (housing, appurtenances, office, retail and industrial) and areas (i.e. central, suburbs, hinterlands), at municipal level, over the sample period running from the second semester 2002 to the second semester 2016. To construct the provincial house prices series for the residential property, I take the average value between the minimum and maximum house prices (for housing category) of the corresponding regional capital. Given the presence of missing data, I discard

Table 3.4: Average pairwise cross-section correlations.

	Northwest					Northeast			
	Sales		HP			Sales		HP	
	Δy_{it}	Res	Δy_{it}	Res	Provinces	Δy_{it}	Res	Δy_{it}	Res
ALESSANDRIA	0.302	0.021	0.141	0.075	BELLUNO	0.256	-0.025	0.286	0.067
AOSTA	0.251	-0.044	0.172	-0.001	BOLOGNA	0.519	-0.008	0.241	0.110
ASTI	0.406	0.041	0.019	0.000	FERRARA	0.494	0.016	0.111	-0.010
BERGAMO	0.271	-0.016	0.092	-0.015	FORLÍ-CESENA	0.425	0.004	0.110	-0.072
BIELLA	0.123	0.000	-0.003	0.005	MODENA	0.484	0.018	0.088	0.026
BRESCIA	0.157	-0.017	0.017	-0.013	PADOVA	0.311	0.002	0.312	0.061
COMO	0.433	0.044	0.093	0.050	PARMA	0.511	0.006	0.230	0.038
CREMONA	0.382	0.045	0.045	0.075	PIACENZA	0.413	-0.061	0.341	0.045
CUNEO	0.197	-0.047	0.070	0.073	PORDENONE	0.385	-0.020	0.349	0.078
GENOVA	0.224	-0.055	-0.098	-0.025	RAVENNA	0.238	-0.081	0.300	0.007
IMPERIA	0.475	0.063	0.181	-0.025	REGGIO EMILIA	0.382	-0.039	0.269	0.083
LA SPEZIA	0.382	0.099	-0.099	-0.112	RIMINI	0.418	0.077	0.389	0.082
LECCO	0.141	0.053	0.054	0.038	ROVIGO	0.312	-0.044	0.105	0.013
LODI	0.280	0.073	0.136	0.010	TREVISO	0.389	-0.029	0.321	0.078
MANTOVA	0.259	-0.038	0.208	0.113	UDINE	0.317	-0.011	0.307	0.103
MILANO	0.313	0.057	0.060	0.037	VENEZIA	0.403	-0.006	0.325	0.069
NOVARA	0.406	0.026	0.107	0.017	VERONA	0.434	0.027	0.373	0.058
PAVIA	0.278	0.040	0.176	0.073	VICENZA	0.390	-0.020	-0.013	-0.007
SAVONA	0.297	0.022	0.223	0.043					
SONDRIO	0.352	0.074	0.066	-0.042					
TORINO	0.487	0.035	0.133	-0.025					
VARESE	0.163	-0.037	0.136	0.010					
VERBANIA	0.227	0.020	0.115	0.030					
	Centre					South			
	Sales		HP			Sales		HP	
	Δy_{it}	Res	Δy_{it}	Res	Provinces	Δy_{it}	Res	Δy_{it}	Res
ANCONA	0.415	0.046	0.445	0.004	AVELLINO	0.219	0.018	0.107	0.042
AREZZO	0.424	0.061	0.393	0.049	BARI	0.234	-0.046	0.410	0.003
ASCOLI PICENO	0.204	0.045	0.358	-0.014	BENEVENTO	0.071	-0.086	0.192	-0.022
FIRENZE	0.356	0.019	0.413	0.031	BRINDISI	0.262	-0.004	0.331	0.096
FROSINONE	0.248	0.097	0.351	0.043	CAMPOBASSO	0.209	0.009	0.278	0.001
GROSSETO	0.238	0.055	0.113	-0.080	CASERTA	0.159	0.087	0.336	0.057
LATINA	0.064	0.024	0.387	0.052	CATANZARO	-0.025	-0.064	0.245	0.048
LIVORNO	0.450	0.081	0.498	0.042	CHIETI	0.032	-0.008	0.279	0.051
LUCCA	0.391	0.015	0.459	-0.013	COSENZA	0.273	0.082	0.294	-0.022
MASSA	0.225	0.044	0.211	-0.039	CROTONE	0.124	-0.002	0.201	-0.005
PERUGIA	0.279	0.085	0.080	-0.073	FOGGIA	0.230	-0.006	0.326	-0.008
PESARO	0.285	-0.072	0.365	-0.017	ISERNIA	0.169	-0.049	0.393	0.164
PISA	0.316	-0.006	0.462	-0.013	LECCE	0.054	0.003	0.118	-0.030
PISTOIA	0.212	-0.032	0.377	0.001	MATERA	0.238	0.087	0.104	0.104
PRATO	0.342	-0.068	0.505	-0.008	NAPOLI	0.166	0.004	0.432	0.104
RIETI	0.440	0.081	0.344	-0.005	PESCARA	0.163	0.021	0.334	0.060
ROMA	0.346	0.085	0.374	-0.015	POTENZA	0.102	-0.054	0.220	0.030
SIENA	0.166	-0.076	0.253	0.009	REGGIO CALABRIA	-0.033	0.004	0.280	0.041
TERNI	0.115	-0.067	0.372	0.050	SALERNO	0.095	0.045	0.299	0.016
VITERBO	0.339	0.002	0.418	0.053	TARANTO	0.306	0.026	0.282	-0.038
					TERAMO	0.065	-0.018	0.162	0.101
					VIBO VALENTIA	0.078	0.027	0.141	0.023
	Islands								
	Sales		HP						
	Δy_{it}	Res	Δy_{it}	Res					
AGRIGENTO	0.207	-0.063	-0.095	-0.138					
CALTANISSETTA	0.201	-0.072	0.297	-0.054					
CATANIA	0.506	0.069	0.352	-0.047					
ENNA	-0.021	-0.197	0.200	-0.039					
MESSINA	0.363	0.122	0.119	0.094					
PALERMO	0.500	-0.020	0.339	-0.242					
RAGUSA	0.276	0.024	0.273	0.081					
SIRACUSA	0.259	-0.069	0.214	-0.014					
TRAPANI	0.365	0.008	0.145	0.029					

Note. Δy_{it} is the variable in log-differences while Res corresponds to residuals of the country-specific $VARX^{*}(1,1)$.

the series for the provinces of L'Aquila and Macerata.[16] I compute the real house prices by applying the Italian Consumer price index (CPI), downloaded from the statistical database of the Italian National Institute of Statistics (ISTAT), on the provincial house prices series.

As for the transaction volumes, I use time series (available at quarterly frequency) for the number of normalized transactions (NNT)[17], collected from the publicly available database of the Real Estate Market Observatory - Italian Revenue Agency ("Agenzia delle Entrate - Osservatorio del Mercato Immobiliare"), covering the $2004Q1 - 2016Q4$ time span. It is important to observe that, in order to match the semi-annual data frequency of house prices, I aggregate the quarterly data on volumes, by taking the sum over two consecutive quarters. Given the lack of volumes data for 11 provinces, the final number of provinces considered is equal to 93 (see Table 3.1).[18]

Given the lack of data for most of provinces which belong to Sardinia, I exclude this region from the analysis.

Since the time series for prices and volumes are not stationary, I apply the first order difference operator to the log transformation of the real house prices and of the number of transactions.[19]

3.4.2 Results

Figures 3.1 and 3.2 show the structural impulse responses of house prices and transaction volumes (in levels) to a negative housing demand shock to 10 Italian regional capitals.[20] Given the use of first order difference of the log transformation of the real house prices and of the transaction volumes, the impulse responses for the series in levels are computed as cumulative sum of ones obtained for the first order difference.

In line with Eickmeier & Ng (2015), all the figures show the bootstrap median estimates (black line) and the 90 percent confidence intervals (shadow area) obtained through the Median Target (MT) approach. Each figure displays two Charts. Chart *a* is the plot

[16] The house prices series for L'Aquila reveal a relevant number of missing entries. Particularly, data for the period July 2009 - June 2012 are not provided. This lack of observations might be due to the heavy earthquakes which devastated part of the Central Italy, including L'Aquila and its neighbourhood zones, on April 2009. I also discard the house prices series for Macerata, where the last observation is missing.

[17] The NNT is the number of "standardized" units sold, taking into account the share of property transferred.

[18] The missing time series series refer to the following provinces: Bolzano/Bozen, Trento, Gorizia and Trieste (Northeast), where the cadastre and/or the land registry are managed by local administrations, and Monza e della Brianza (Northwest), Fermo (Centre), Barletta-Andria-Trani (South), Carbonia-Iglesias, Medio-Campidano, Ogliastra and Olbia-Tempio (Islands).

[19] Results based on the autocorrelation functions (ACF) plots are available upon request.

[20] Given the focus on house prices and transaction volumes of regional capitals in each province, "regional capital" and "province" are used, in this section, as synonymous.

Table 3.5: GVAR models, regional capitals and neighbours.

GVARs	Regional capitals	Neighbours
Northwest	Torino	Alessandria, Aosta, Asti, Biella, Cuneo and Vercelli
	Genova	Alessandria, La Spezia and Savona
	Milano	Bergamo, Cremona, Lodi, Novara, Pavia and Varese
Northeast	Venezia	Padova, Pordenone, Rovigo, Treviso and Udine
	Bologna	Ferrara, Modena and Ravenna
Centre	Firenze	Arezzo, Lucca, Pisa, Pistoia, Prato and Siena
	Roma	Frosinone, Latina, Rieti and Viterbo
South	Napoli	Avellino, Benevento, Caserta and Salerno
	Bari	Brindisi, Matera, Potenza and Taranto
Islands	Palermo	Agrigento, Caltanissetta, Enna, Messina and Trapani

Note. For each regional capital, the corresponding neighbours are identified through a contiguity-based method. According to this criteria, it is possible to define as neighbours those provinces (regional capitals) which share a common border.

of the impulse response of the m-th main regional capital house prices and sales to a negative housing demand shock occurring to the m-th main regional capital. Chart b is the plot of the impulse response of the house prices and sales of the m-th main regional capital's neighbours to a negative housing demand shock arising from the m-th main regional capital.

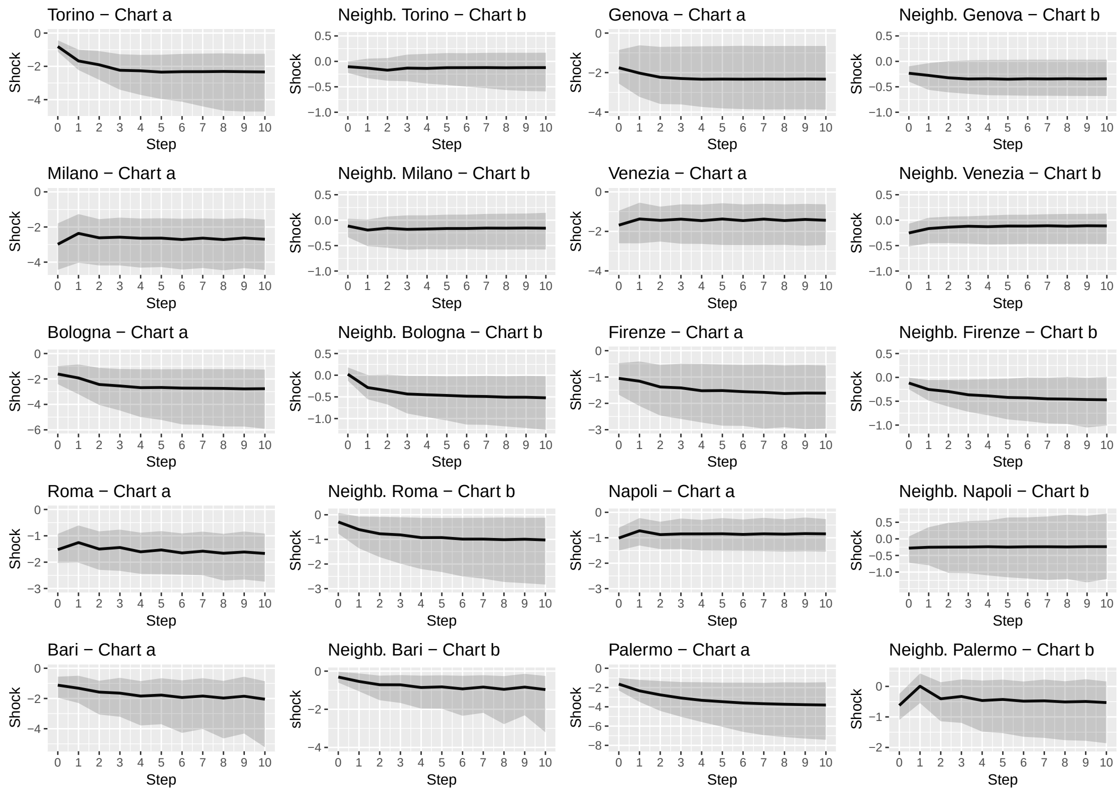
For the sake of simplicity, I define the response of the m-th main regional capital's aggregate to the exogenous shock arising from the m-th main regional capital as "Domestic response", while the response of the neighbours' house prices and volumes to the exogenous shock occurring to the m-th main regional capital is labelled as "Spillover effect". In Chart b, the spatial dimension is captured by considering the provinces which share a common border with the main province (say, its neighbours). Since a regional capital is likely to share common borders with more than one province (see Table 3.5), I aggregate the impulse responses of individual neighbours using their value added reported on 2014 as weights (see Eickmeier & Ng, 2015; Vansteenkiste & Hiebert, 2011).[21]

The orthogonalized impulse response are to a one standard deviation negative shock to housing demand and they are computed over a 10 semesters (e.g. 5 years) forecast horizon.

All the impulse responses of house prices in regional capitals ("domestic response") are negative and statistically significant (see Figure 3.1, Chart a). Inspection of Figure 3.1

[21]To weight impulse responses of individual units, Eickmeier & Ng (2015) use the PPP-adjusted GDP averaged over 2006 – 2008, while Vansteenkiste & Hiebert (2011) use 2007 real GDP to aggregate the impulse responses of a group of Euro area countries.
In my analysis, I use provincial value added downloaded from the Statistical Database of the Italian National Institute of Statistic (ISTAT). The latest available observations refer to 2014.

Figure 3.1: Responses of the real house prices level in main regional capitals and neighbours to a negative housing demand shock occurring in main regional capitals.



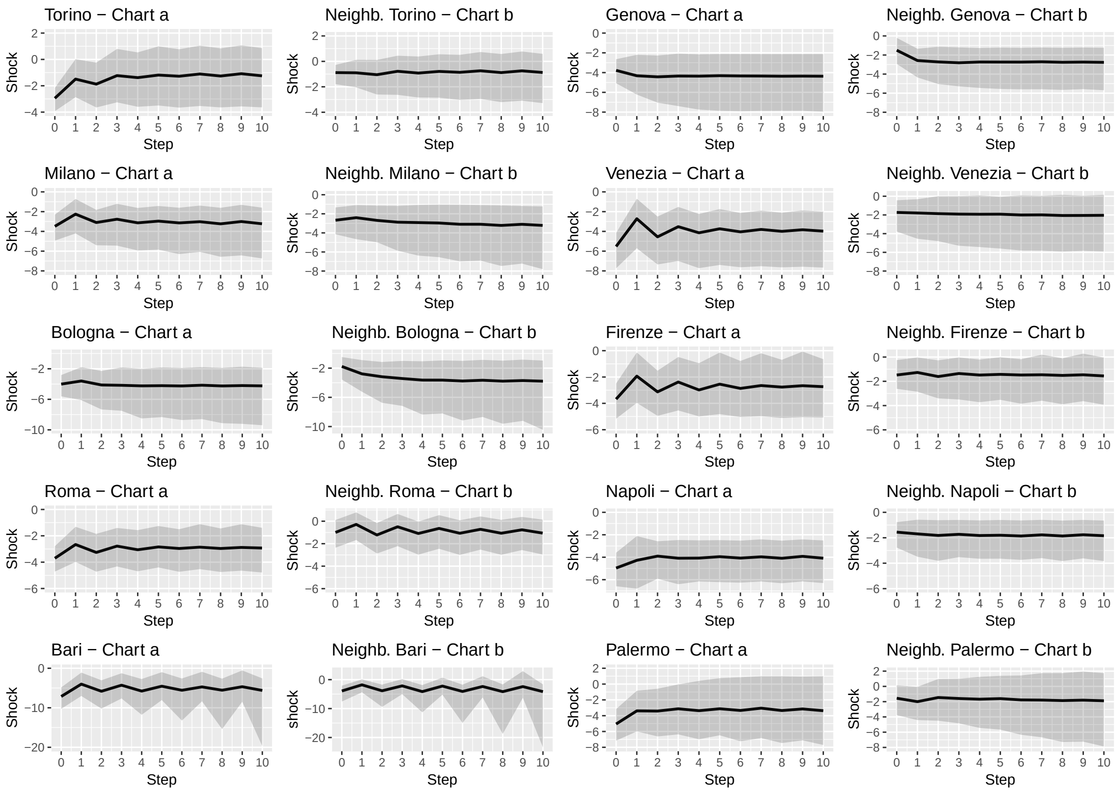
Note. Impulse responses of real house prices level to a one standard deviation shock (in percentage) occurring in the Italian main regional capitals. The Bootstrap median estimates (black line) and the 90 percent confidence intervals (shadow area) are reported. Chart a shows the median response of main regional capitals' real house prices to domestic shock. Chart b presents the median response of neighbours house prices to shocks arising from the corresponding main regional capital housing market.

(Chart a) shows that the largest "domestic response" of the house prices level, on impact, is recorded for Milano (2.98 percent) and the lowest is for Torino (0.81 percent). Figure 3.1 (Chart a) shows that the "domestic" negative response persists and it converges to a new equilibrium value, at most, over a five-year horizon. This finding is confirmed by Table 3.6 (panel a) showing the *Within domestic ratio*, that is the "domestic response" for each forecast horizon relative to the one occurring at time 0. Table 3.6 (panel a) shows that the index slightly increases reaching the highest value in the last semesters, in almost all main regional capitals.

This results are also confirmed by inspecting Figure 3.3, which shows the "domestic response" of house prices changes (Chart *a*) and the corresponding "spillover effect" (Chart *b*) to a negative housing demand shock to 10 Italian regional capitals. As shown in Figure 3.3 (Chart *a*), changes in house prices in response to a negative housing demand shock to main regional capitals become smaller as the forecast horizons increase.

The analysis of the "ripple effect" is carried out by, first, inspecting the "spillover effect" on impact. In order to interpret the empirical evidence shown in Chart *a* and Chart *b*,

Figure 3.2: Responses of the transaction volumes level in main regional capitals and neighbours to a negative housing demand shock occurring in main regional capitals.



Note. Impulse responses of transaction volumes level to a one standard deviation shock (in percentage) occurring in the Italian main regional capitals. The Bootstrap median estimates (black line) and the 90 percent confidence intervals (shadow area) are reported. Chart a shows the median response of main regional capitals' transaction volumes to domestic shock. Chart b presents the median response of neighbours transaction volumes to shocks arising from the corresponding main regional capital housing market.

I compute the *Spillover index* which is measured, on impact, by the ratio of the median response (at horizon 0) of the neighbours (the "spillover effect") to the median response (at horizon 0) of the main regional capitals (the "domestic response"). Table 3.7 (panel a) shows that, on impact, the largest transmission of the shock to neighbours is recorded in the main cities of Mezzogiorno, such as Palermo (37.95 percent), Bari (27.74 percent) and Napoli (27.56 percent).[22] The lowest values of the transmission mechanism on impact are recorded in Bologna (−1.40 percent) and Milano (3.84 percent). The relative small values of the impact *Spillover index* in Torino, Genova, Venezia and Firenze are 12.91, 13.29, 14.95 and 11.24 percent, respectively.

Moreover, in line with previous empirical studies on "ripple effect", I need to compare the plots of Chart *a* and Chart *b* by computing a *Spillover index* for horizons beyond time 0. For this purpose, I choose to focus on a time span involving at most five years

[22]ISTAT defines Mezzogiorno as the macro-area which includes the six Southern regions and the Islands of Sardinia and Sicily

Table 3.6: Within domestic ratio.

Panel a: Within domestic ratio for real house prices level.

horizon (h)	Torino	Genova	Milano	Venezia	Bologna	Firenze	Roma	Napoli	Bari	Palermo
h=0 / h=0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
h=1 / h=0	2.071	1.154	0.792	0.814	1.192	1.102	0.824	0.717	1.180	1.427
h=2 / h=0	2.355	1.273	0.877	0.856	1.509	1.310	0.985	0.867	1.410	1.680
h=3 / h=0	2.762	1.309	0.865	0.820	1.580	1.344	0.947	0.839	1.471	1.877
h=4 / h=0	2.802	1.330	0.884	0.865	1.664	1.447	1.054	0.838	1.643	2.024
h=5 / h=0	2.899	1.326	0.882	0.815	1.656	1.441	1.007	0.834	1.586	2.110
h=6 / h=0	2.868	1.329	0.910	0.863	1.685	1.479	1.082	0.860	1.728	2.193
h=7 / h=0	2.866	1.326	0.883	0.819	1.692	1.505	1.038	0.835	1.643	2.239
h=8 / h=0	2.850	1.328	0.911	0.861	1.701	1.548	1.090	0.849	1.761	2.274
h=9 / h=0	2.870	1.324	0.880	0.829	1.723	1.531	1.059	0.828	1.653	2.300
h=10 / h=0	2.885	1.328	0.904	0.854	1.714	1.534	1.095	0.837	1.829	2.321

Panel b: Within domestic ratio for transaction volumes level.

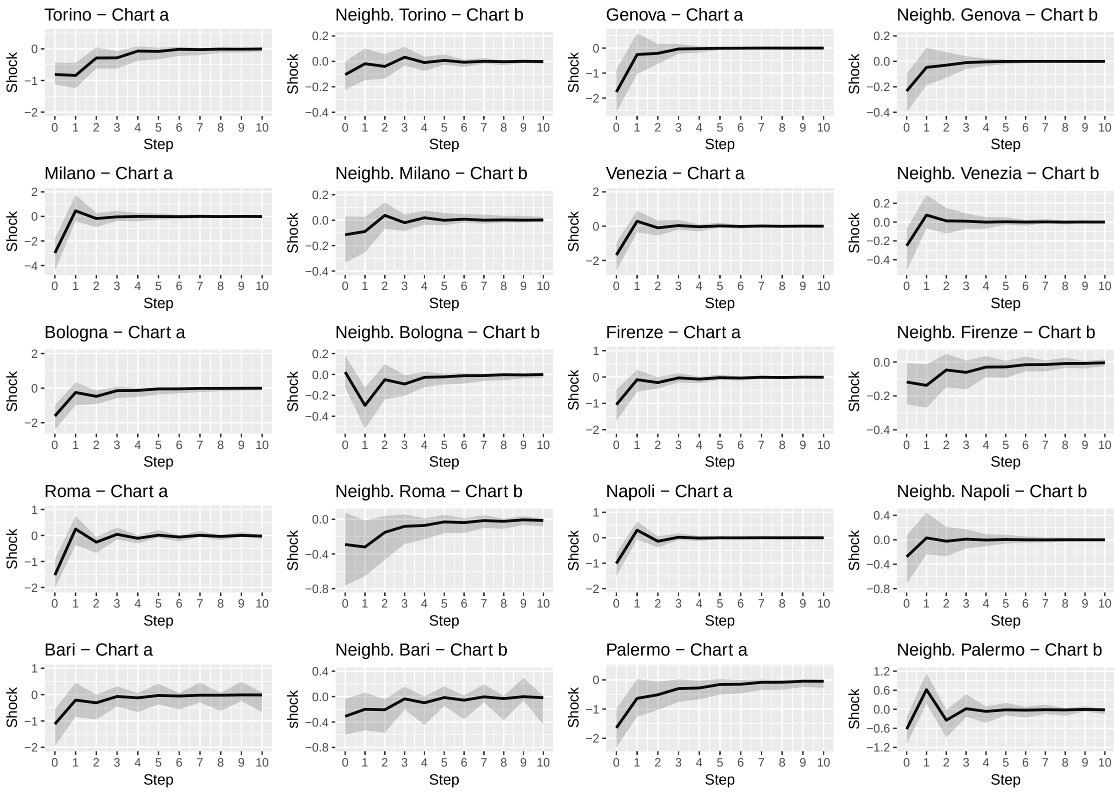
horizon (h)	Torino	Genova	Milano	Venezia	Bologna	Firenze	Roma	Napoli	Bari	Palermo
h=0 / h=0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
h=1 / h=0	0.507	1.144	0.646	0.492	0.898	0.525	0.718	0.861	0.561	0.671
h=2 / h=0	0.636	1.173	0.889	0.824	1.028	0.850	0.881	0.786	0.816	0.675
h=3 / h=0	0.417	1.151	0.793	0.637	1.039	0.647	0.751	0.823	0.600	0.619
h=4 / h=0	0.469	1.154	0.898	0.749	1.057	0.810	0.826	0.821	0.809	0.666
h=5 / h=0	0.403	1.142	0.849	0.676	1.050	0.690	0.767	0.796	0.636	0.618
h=6 / h=0	0.433	1.149	0.900	0.732	1.059	0.778	0.800	0.821	0.780	0.660
h=7 / h=0	0.375	1.151	0.865	0.688	1.036	0.719	0.773	0.798	0.661	0.603
h=8 / h=0	0.427	1.156	0.927	0.723	1.059	0.751	0.800	0.824	0.778	0.663
h=9 / h=0	0.368	1.153	0.861	0.693	1.048	0.723	0.778	0.790	0.654	0.622
h=10 / h=0	0.423	1.156	0.925	0.719	1.056	0.743	0.790	0.823	0.786	0.667

Note. The *Within domestic ratio* is computed as the ratio between the median impulse response of main regional capitals house prices (transaction volumes) to domestic shock at each h-step ahead and the median impulse response of main regional capitals house prices (transaction volumes) to domestic shock at time 0.

ahead (ten semesters). Both the numerator and the denominator of the *Spillover index* for the different forecast horizons are responses to a 1 standard deviation negative housing demand shock to the regional capital occurring at time 0. I focus, first, on discussing results for the house prices spillovers. From Table 3.7 (panel a), it can be observed that the *Spillover index* decreases (over a time span involving forecast horizons beyond time 0 up to the next five years) in three Northern cities, such as Torino (from 7.92 to 5.23 percent), Milano (from 8.21 to 5.85 percent) and Venezia (from 12.04 to 7.82 percent), and in two Mezzogiorno cities, such as Napoli (from 34.84 to 27.79 percent) and Palermo (the average *Spillover index*, across forecast horizons beyond time 0 is equal to 11.84 percent, lower than the initial impact equal to 37.95 percent). All the remaining cities exhibit an heterogeneous increase in the *Spillover index*. More specifically, a moderate increase can be observed in the Northern cities: Genova, from 13.61 to 14.64 percent, and Bologna, from 14.77 to 18.88 percent. The cities in Central Italy and Bari exhibit the largest increase (Firenze, from 21.95 to 29.18 percent, Roma, from 47.84 to 61.27 percent, and Bari, from 41.02 to 47.22 percent).

The convergence of the *Spillover index* to an equilibrium value is fastest in Northern

Figure 3.3: Responses of the house price changes in main regional capitals and neighbours to a housing demand shock occurring in main regional capitals.



Note. Impulse responses of real house prices changes to a one standard deviation shock (in percentage) occurring in the Italian main regional capitals. The Bootstrap median estimates (black line) and the 90 percent confidence intervals (shadow area) are reported. Chart a shows the median response of main regional capitals' real house prices changes to domestic shock. Chart b presents the median response of neighbours house prices changes to shocks arising from the corresponding main regional capital housing market.

cities, such as Genova, Milano and Bologna, where the index reaches an equilibrium value over a one-year horizon, while it takes longer, say $2-3$ years, in the two Central regions, Firenze and Roma, where the *Spillover index* reaches equilibrium values equal to around 28 and 61 percent, respectively, and in two Mezzogiorno cities, such as Napoli and Palermo, in which the *Spillover index* reaches values equal to around 28 and 13 percent, respectively.

I now turn the focus on the responses of transaction volumes to negative housing demand shock. From Figure 3.2, it can be observed that, on impact, both the "domestic response" and the "spillover effect" are larger than the ones recorded in house prices. Similarly to the results obtained for house prices level, all the "domestic responses" are negative and statistically significant, with the exception of Torino and Palermo where the median impulse response becomes not statistically significant for long forecast horizons (see Figure 3.2, Chart a). The largest impact "domestic response" of the transaction

Table 3.7: Spillover index (in percentage).

Panel a: Spillover index for real house prices level.

horizon (h)	Torino	Genova	Milano	Venezia	Bologna	Firenze	Roma	Napoli	Bari	Palermo
h=0	12.912	13.287	3.844	14.947	-1.404	11.243	19.102	27.561	27.743	37.946
h=1	7.923	13.606	8.207	12.036	14.772	21.952	47.836	34.843	41.024	-0.232
h=2	9.081	14.431	6.043	9.516	14.647	21.534	51.292	28.281	45.220	14.716
h=3	5.906	15.034	6.992	8.647	17.020	25.854	56.491	28.974	43.553	10.692
h=4	6.088	14.632	6.587	8.784	16.834	25.501	57.632	27.989	46.635	13.943
h=5	5.304	15.073	6.231	8.407	17.485	27.752	60.270	29.355	46.471	12.372
h=6	5.325	14.687	6.029	8.020	17.812	27.639	59.856	27.484	48.119	13.442
h=7	5.268	14.840	5.923	7.975	17.993	28.469	62.325	28.156	45.435	12.863
h=8	5.461	14.642	5.805	8.166	18.485	27.997	60.858	28.269	48.388	13.635
h=9	5.299	14.865	5.891	7.845	18.293	28.888	61.476	28.082	45.201	13.088
h=10	5.230	14.644	5.853	7.822	18.880	29.177	61.270	27.794	47.217	13.908

Panel b: Spillover index for transaction volumes level.

horizon (h)	Torino	Genova	Milano	Venezia	Bologna	Firenze	Roma	Napoli	Bari	Palermo
h=0	29.963	39.705	76.593	31.330	44.182	40.196	26.957	31.371	54.905	30.716
h=1	60.225	59.733	106.715	65.939	77.363	65.745	10.722	39.653	44.632	59.034
h=2	55.864	61.421	86.677	40.909	77.186	51.363	37.569	46.568	66.294	42.703
h=3	63.437	64.711	103.948	54.267	82.224	56.969	17.791	42.296	50.230	50.775
h=4	66.546	62.680	93.177	46.545	85.863	49.755	35.701	44.759	71.574	49.845
h=5	66.461	63.442	100.071	51.377	86.488	56.169	22.549	45.609	49.019	51.273
h=6	67.431	63.148	98.572	49.533	88.583	51.656	36.190	45.660	73.410	53.091
h=7	66.940	62.332	102.543	52.390	88.105	55.118	25.151	44.649	50.206	58.812
h=8	70.035	63.215	99.423	51.620	89.110	54.765	36.084	45.445	74.749	55.575
h=9	68.560	62.874	103.223	53.574	88.155	54.980	26.543	44.950	51.534	57.363
h=10	70.065	63.408	99.578	51.288	89.451	56.943	36.135	45.022	73.955	55.955

Note. The *spillover index* is computed as the ratio between the median impulse response of main regional capitals house prices (transaction volumes) to domestic shock and the median response of neighbours house prices (transaction volumes) to shock arising from main regional capitals housing market, at h-step ahead.

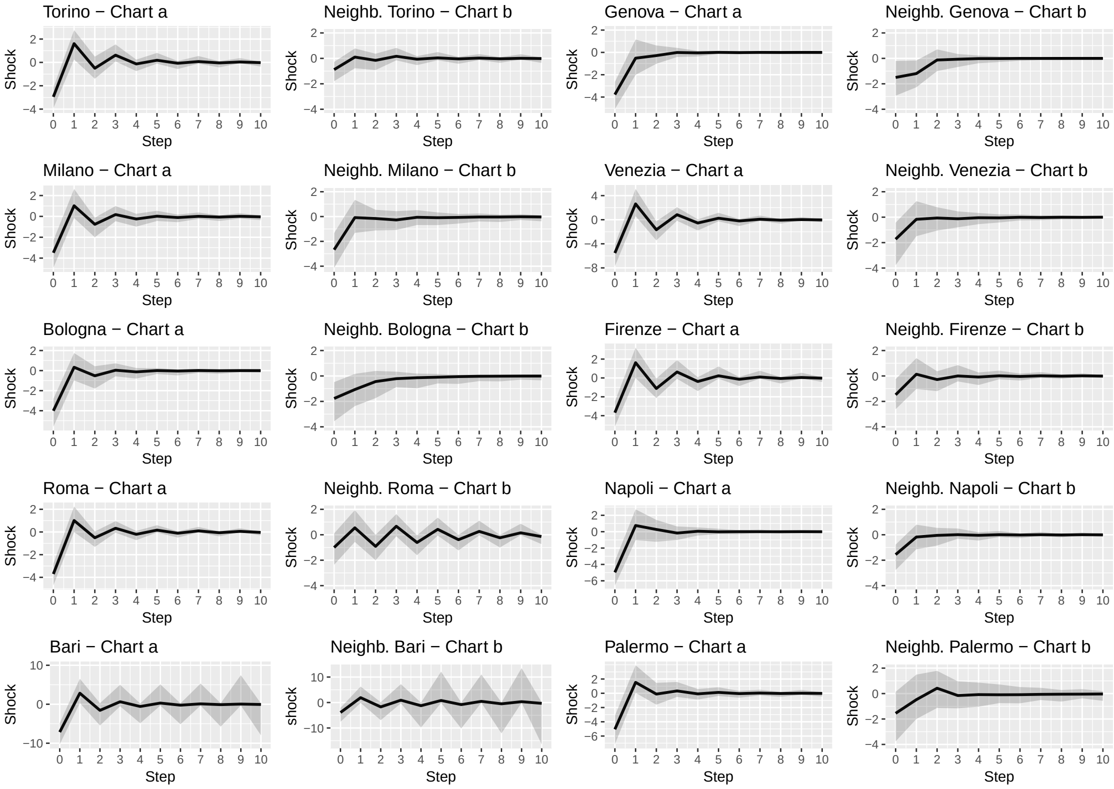
volumes level is recorded in Bari (7.14 percent), while the lowest is associated once more in Torino (2.95 percent).

Differently from the results obtained for house prices, the "domestic response" of transaction volumes level reaches its maximum value over one-year horizon, in almost all main cities. As shown in Table 3.6 (panel b), the *Within domestic ratio* reaches its peak throughout two semesters in almost all cities (with the exception of Milano and Bologna) showing values of the ratio larger than the ones recorded at a five-year horizon.

Figure 3.4 shows the impulse response of main regional capital's transaction volumes changes (Chart a) and neighbours' transaction volume changes (Chart b) to a negative housing demand shock to 10 Italian regional capitals. If I focus on the "domestic response", it can be seen from Figure 3.4 (Chart a) that the transaction volumes changes strongly react to the exogenous shock over the first semester before reaching their base value.

To investigate the presence of a "ripple effect" in the Italian main provinces, I also focus on the *Spillover index* constructed for transaction volumes (see Table 3.7, panel b). At

Figure 3.4: Responses of the transaction volumes changes in main regional capitals and neighbours to a housing demand shock occurring in main regional capitals.



Note. Impulse responses of transaction volumes changes to a one standard deviation shock (in percentage) occurring in the Italian main regional capitals. The Bootstrap median estimates (black line) and the 90 percent confidence intervals (shadow area) are reported. Chart a shows the median response of main regional capitals' transaction volumes changes to domestic shock. Chart b presents the median response of neighbours transaction volumes changes to shocks arising from the corresponding main regional capital housing market.

horizon 0, the largest transmission of the housing demand shock to neighbours transaction volumes level is observed for Milano (76.59 percent) and, to less extent, in Bari (54.91 percent), while the *Spillover index* is similar for the other main cities, with values ranging from 26.96 percent (Roma) to 44.18 percent (Bologna). Moreover, I focus on the transitional path of the volumes "spillover effect" from time 0 to a five-year forecast horizon, by comparing the *Spillover index* corresponding to a forecast horizon beyond time 0 with the one associated with a five-year horizon. It can be seen from Table 3.7 (panel b) that the *Spillover index* strongly increases in Bari (from 44.63 to 73.96 percent) and in Roma (from 10.72 to 36.14 percent). Since confidence intervals for Bari get dramatically wider as the forecast horizons increase (see Figure 3.2), I focus only on Bari spillover effect over a short-run forecast horizon. More specifically the empirical results for Bari suggest an average *Spillover index* beyond time 0 up to e.g. 2 years equal to 58 percent, that is a value bigger than the one for the impact effect. The other main city in Mezzogiorno, Napoli, shows an increase in the *Spillover index*, since there

is a moderate increase by 5 percent. A rise of the *Spillover index* across forecast horizon beyond time 0 is also recorded by the Northern cities, including: Bologna (from 77.36 to 89.45 percent), Torino (from 60.23 to 70.07 percent) and, to less extent, Genova (from 59.73 to 63.41 percent). However, all the three Northern cities report values of the index at a five-year horizon decisively larger than the ones reported at time 0 (44.18, 29.96 and 39.71 percent, respectively). All the remaining main cities exhibit a decrease of the *Spillover index*: Venezia (from 65.94 to 51.29), Firenze (from 65.75 to 56.94), Milano (from 106.72 to 99.58 percent) and, to less extent, Palermo (from 59.03 to 55.96 percent). However, the average *Spillover index*, across forecast horizon beyond time 0, in each of these four cities is larger than the index measured at time 0 (the average values are equal to 51.74, 55.34, 99.39 and 53.44 percent, respectively).
The convergence of the volumes *Spillover index* to its equilibrium value is slower than the one observed for the house prices. The fastest convergence (over two years) is recorded in two Northern cities, such as Genova and Bologna, and only in one city of the Mezzogiorno, Napoli. All the remaining cities show a slower convergence process, taking the whole five-year horizon.

To summarize, the structural impulse response (IRF) analysis together with the associated *Spillover index* provides some interesting findings. First, contrary to a large body of literature, this study does not find evidence of a "ripple effect" in house prices. There is evidence of neighbours small response to a negative housing demand shock to the main regional capital, especially in the North of Italy. The only exception is Roma, where the *Spillover index* increases over the whole forecast period, showing a "spillover effect" at five-year horizon three times bigger than the one reported on impact.
I find that transaction volumes largely spill over across regional capitals and neighbours in response to the negative housing demand shock. In all the 10 main regional capitals, the *Spillover index* at five-year horizon is larger than its value on impact. My findings are consistent with the study of Tsai (2014), which focuses on UK housing market. In particular, the empirical evidence in this chapter supports the presence of a "ripple effect" in transaction volume. My findings are consistent with a number of studies which focus on the impact of unobserved shocks to fundamentals on price-volume correlation. While the literature concentrates on reduced form shocks to fundamentals, I focus on the response to an unobserved structural form shock to fundamentals, interpreted as negative housing demand shock. Focusing on the "domestic response", my findings show a stronger reaction, on impact, of transaction volumes than house prices. This results support those of Hort (2000) and Clayton *et al.* (2010), which find a reaction of the number of sales and the turnover rates (respectively) to reduced form shocks to fundamentals, on impact, larger than the response of house prices. The two housing market aggregates show a different behaviour beyond time 0. In line with the study

of Andrew & Meen (2003), which focuses on the response of house prices and transaction volumes (changes) to an interest rate shock in UK housing market, I find that house prices slightly decrease over the whole forecast period, while transaction volumes strongly react over few semesters, say $2-3$ semesters, before reaching their base value, in almost all the 10 main regional capitals. Finally, I find evidence of an heterogeneity in the ripple effect given a different propagation of the negative housing demand shock arising in each dominant unit to the price and the volumes of neighbours.

3.5 Conclusions

In this chapter, I have contributed to the literature on the spatio-temporal diffusion house prices, which is known as "ripple effect". First, I have focused not only on house prices but also on transaction volumes. The bi-annual dataset is for 93 Italian provinces, over the period $2004-2016$. Second, I have explored heterogeneity in the "ripple effect" by considering different dominant units. Third, I have also contributed to the literature on price-volume co-movement associated to reduced form shocks to the fundamentals, by focusing on a structural form innovation identified as negative housing demand shock. The use of a structural shock allows to circumvent the issue related to the lack of provincial data for fundamental drivers of house prices such as interest rates on loan and income.

The spillover analysis has been carried out by using a GVAR model based on a spatial exogenous regressor obtained from the construction of a spatial weight matrix (spatial econometric approach). The structural housing demand shocks in each of the 10 Italian main regional capitals have been identified by using theory-driven sign restrictions.

The structural impulse response functions obtained from the estimated GVAR allow to address the three aforementioned issues.

As for the analysis on "ripple effect", I do not find evidence of a strong propagation mechanism of the housing demand shocks on neighbours house prices. Oppositely, in line with the study of Tsai (2014) which finds evidence of a "ripple effect" in transaction volumes for UK, I find a significant transmission mechanism of the exogenous shock to neighbours through sales, in almost all the 10 provinces under investigation.

Second, there is evidence of heterogeneity in the ripple effect given the different responses to a shock to each dominant unit.

Finally, I also focus on the relationship between house prices and transaction volumes in response to a housing demand shock. This results support those of Hort (2000) and Clayton *et al.* (2010), which find a reaction of the number of sales and the turnover rates (respectively) to reduced form shocks to fundamentals, on impact, larger than the response of house prices. The two housing market aggregates show a different behaviour beyond time 0. In line with the study of Andrew & Meen (2003), which focuses on the response of house prices and transaction volumes (changes) to an interest rate shock in UK housing market, I find that house prices slightly decrease over the whole forecast period, while transaction volumes strongly react over few semesters, say $2-3$ semesters, before reaching their base value, in almost all the 10 main regional capitals.

The evidence of heterogeneity in the ripple effect implies the existence of segmented housing markets regardless of the geographical location and it might suggest housing market policy intervention tailored to the local condition of a given housing market.

Bibliography

Andrew, Mark, & Meen, Geoffrey. 2003. House price appreciation, transactions and structural change in the British housing market: a macroeconomic perspective. *Real Estate Economics*, **31**(1), 99–116.

Arias, Jonas, Rubio-Ramirez, Juan Francisco, & Waggoner, Daniel F. 2014. *Inference based on svar identified with sign and zero restrictions: Theory and applications.* International Finance Discussion Papers 1100. Board of Governors of the Federal Reserve System.

Berkovec, James A, & Goodman, John L. 1996. Turnover as a measure of demand for existing homes. *Real Estate Economics*, **24**(4), 421–440.

Blanchard, Olivier Jean, & Quah, Danny. 1989. The Dynamic Effects Of Aggregate Demand And Supply Disturba. *The American Economic Review*, **79**(4), 655–673.

Brady, Ryan R. 2011. Measuring the diffusion of housing prices across space and over time. *Journal of Applied Econometrics*, **26**(2), 213–231.

Brady, Ryan R. 2014. The spatial diffusion of regional housing prices across US states. *Regional Science and Urban Economics*, **46**, 150–166.

Bühlmann, Peter. 1997. Sieve bootstrap for time series. *Bernoulli*, **3**(2), 123–148.

Cesa-Bianchi, Ambrogio. 2013. Housing cycles and macroeconomic fluctuations: A global perspective. *Journal of International Money and Finance*, **37**, 215–238.

Cesa-Bianchi, Ambrogio, Pesaran, M Hashem, & Rebucci, Alessandro. 2014. *Uncertainty and economic activity: A global perspective.* CESifo Working Paper Series 4736. CESifo Group Munich.

Chudik, Alexander, & Fidora, Michael. 2011. *Using the global dimension to identify shocks with sign restrictions.* Working Paper No. 1318. European Central Bank.

Chudik, Alexander, & Pesaran, M Hashem. 2016. Theory and practice of GVAR modelling. *Journal of Economic Surveys*, **30**(1), 165–197.

Clayton, Jim, Miller, Norman, & Peng, Liang. 2010. Price-volume correlation in the housing market: causality and co-movements. *The Journal of Real Estate Finance and Economics*, **40**(1), 14–40.

Conley, Timothy G., & Topa, Giorgio. 2002. Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics*, **17**(4), 303–327.

De Wit, Erik R, Englund, Peter, & Francke, Marc K. 2013. Price and transaction volume in the Dutch housing market. *Regional Science and Urban Economics*, **43**(2), 220–241.

Dees, Stephane, Mauro, Filippo di, Pesaran, M Hashem, & Smith, L Vanessa. 2007a. Exploring the international linkages of the euro area: a global VAR analysis. *Journal of applied econometrics*, **22**(1), 1–38.

Dees, Stephane, Holly, Sean, Pesaran, M Hashem, & Smith, L Vanessa. 2007b. *Long run macroeconomic relations in the global economy*. Working Paper No. 750. European Central Bank.

Diebold, Francis X, & Yilmaz, Kamil. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, **119**(534), 158–171.

Eickmeier, Sandra, & Ng, Tim. 2015. How do US credit supply shocks propagate internationally? A GVAR approach. *European Economic Review*, **74**, 128–145.

Faust, Jon. 1998. The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy*, **49**, 207–244.

Follain, James R, & Velz, Orawin T. 1995. Incorporating the number of existing home sales into a structural model of the market for owner-occupied housing. *Journal of Housing Economics*, **4**(2), 93–117.

Fry, Renee, & Pagan, Adrian. 2007. *Some issues in using sign restrictions for identifying structural VARs*. Working Paper No. 14. National Centre for Econometric Research.

Galesi, Alessandro, & Sgherri, Silvia. 2009. *Regional financial spillovers across Europe: a Global VAR analysis*. Working Paper No. 09/23. International Monetary Fund.

Galesi, Alessandro, & Smith, Vanessa L. 2014. *GVAR Toolbox 2.0*. Available at `https://sites.google.com/site/gvarmodelling/gvar-toolbox`.

Gong, Yunlong, Hu, Jinxing, & Boelhouwer, Peter J. 2016. Spatial interrelations of Chinese housing markets: Spatial causality, convergence and diffusion. *Regional Science and Urban Economics*, **59**, 103–117.

Gray, David. 2012. District house price movements in England and Wales 1997–2007: An exploratory spatial data analysis approach. *Urban Studies*, **49**(7), 1411–1434.

Holly, Sean, Pesaran, M Hashem, & Yamagata, Takashi. 2011. The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics*, **69**(1), 2–23.

Hort, Katinka. 2000. Prices and turnover in the market for owner-occupied homes. *Regional Science and Urban Economics*, **30**(1), 99–119.

Hviid, Simon Juul. 2017. *A regional model of the Danish housing market.* Working Paper No. 121. Danmarks Nationalbank.

Im, Kyung So, Pesaran, M Hashem, & Shin, Yongcheol. 2003. Testing for unit roots in heterogeneous panels. *Journal of econometrics*, **115**(1), 53–74.

Jordà, Òscar. 2005. Estimation and inference of impulse responses by local projections. *American economic review*, **95**(1), 161–182.

Kilian, Lutz, & Lütkepohl, Helmut. 2017. *Structural vector autoregressive analysis.* Cambridge University Press.

Koop, Gary, Pesaran, M Hashem, & Potter, Simon M. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, **74**(1), 119–147.

Krainer, John. 2001. A theory of liquidity in residential real estate markets. *Journal of urban Economics*, **49**(1), 32–53.

Lanne, Markku, & Lütkepohl, Helmut. 2008. Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, **40**(6), 1131–1149.

Meen, Geoffrey. 1999. Regional house prices and the ripple effect: a new interpretation. *Housing studies*, **14**(6), 733–753.

Montagnoli, Alberto, & Nagayasu, Jun. 2015. UK house price convergence clubs and spillovers. *Journal of Housing Economics*, **30**, 50–58.

Pesaran, M Hashem, Schuermann, Til, & Weiner, Scott M. 2004. Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics*, **22**(2), 129–162.

Pijnenburg, Katharina. 2017. The spatial dimension of US house prices. *Urban Studies*, **54**(2), 466–481.

Rigobon, Roberto. 2003. Identification through heteroskedasticity. *Review of Economics and Statistics*, **85**(4), 777–792.

Rubio-Ramirez, Juan F, Waggoner, Daniel F, & Zha, Tao. 2010. Structural vector autoregressions: Theory of identification and algorithms for inference. *The Review of Economic Studies*, **77**(2), 665–696.

Stein, Jeremy C. 1995. Prices and trading volume in the housing market: A model with down-payment effects. *The Quarterly Journal of Economics*, **110**(2), 379–406.

Stewart, Gilbert W. 1980. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM Journal on Numerical Analysis*, **17**(3), 403–409.

Tsai, I-Chun. 2014. Ripple effect in house prices and trading volume in the UK housing market: New viewpoint and evidence. *Economic Modelling*, **40**, 68–75.

Uhlig, Harald. 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, **52**(2), 381–419.

Van Dijk, Bram, Franses, Philip Hans, Paap, Richard, & Van Dijk, Dick. 2011. Modelling regional house prices. *Applied Economics*, **43**(17), 2097–2110.

Vansteenkiste, Isabel. 2007. *Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting.* Working Paper No. 708. European Central Bank.

Vansteenkiste, Isabel, & Hiebert, Paul. 2011. Do house price developments spillover across euro area countries? Evidence from a global VAR. *Journal of Housing Economics*, **20**(4), 299–314.