IMT Institute for Advanced Studies, Lucca

Lucca, Italy

Essays on Empirical Banking

PhD Program in Economics, Markets, Institutions

XX Cycle

By

Caterina Giannetti

2008

The dissertation of Caterina Giannetti is approved.

Program Coordinator: Prof. Fabio Pammolli, Università di Firenze

Supervisor: Prof. Giampiero M. Gallo, Università di Firenze

Tutor: Prof. Fabio Pammolli, Università di Firenze

The dissertation of Caterina Giannetti has been reviewed by:

Prof. Elena Carletti, University of Frankfurt

Prof. Hans Degryse, University of Tilburg

IMT Institute for Advanced Studies, Lucca

2008

To my family
To Eugenio Paladini, Santo Perrotta and Franca Pompilio

Contents

Acknowledgements

1. I am deeply indebted with Giancarlo Spagnolo, Giampiero Gallo, Elena Carletti, Hans Degryse, and Luigi Buzzacchi. I also wish to thank Marcello Messori, Giorgio Rodano, Giuseppe de Arcangelis, Mark Dincecco, Stefano Caiazza, people at CENTER and TILEC (University of Tilburg), as well as people at CORE (Université Catholique de Louvain). I also aknowledge financial support from the FLAFR (Fondazione Lucchese per lAlta Formazione e la Ricerca).

2. Special thanks to Barbara Aretino, Maria Bigoni, Serena Brandini, Elena d'Alfonso, Marilena Filippelli, Manuela Forte, Giulia Gobbo, Marianna Madìa, Luigi Moretti, Valerio Novembre, Carlo Stagnaro, Alessia Savoldi.

Vita

31/07/1978	Born, Roma, Italy
2008	Visiting PhD student Universiteit Van Tilburg - The Netherlands
2006	Visiting PhD student Université Catholique de Louvain - Belgium
2004	Master Degree in Development and International Cooperation. University of Bologna
2002	Degree in Economics. University of Tor Vergata. Final mark: summa cum laude
2005	CIDE Summer School in Microeconometrics
2006	SEEC Summer School in Applied Economics CIDE Summer School in Time Series ESSID Summer School in Industrial Dynamics
2006	CORE - Center For Operations Research and Econometrics - Marie Curie Visiting Fellow - EDNET Program - "Network of the European doctoral program in quantitative economics"
2007	EUROSTAT (Luxembourg) - Paid Trainee
2005	Italian Competition Authority - Unpaid Trainee
2005	'Ca Foscari University (Venice - Italy) - Teaching Assistant Political Economy

Publications

1. Bagella, M., and Caiazza, S., and Carletti, E. and Giannetti, C. and Spagnolo, G. (2008) "Analasi della Concorrenza nei Servizi Bancari Europei: Il punto di vista delle Autorità sulla Concorrenza", Rapporto sul sistema finanziario *forthcoming*
2. C. Giannetti (2007) 'Intensity of Competition and Market Structure in the Italian banking Industry' *CORE Discussion Paper*, N. 41.

Presentations

1. TILEC - Tilburg University - 2008 - Tilburg - The Netherlands
2. DTE - Università La Sapienza -2007 - Rome - Italy
3. ISAE Monitoring Italy 2007 - Rome - Italy
4. EARIE CONFERENCE 2007 - Valencia - Spain
5. CES - Université de Paris I 2007- France
6. ESSID Workshop 2006 - Cargèse - France
7. Doctoral Workshop in Economics 2006 - Louvain-la-Neuve - Belgium
8. ZEW - 2005 - Mannheim - Germany

Abstract

The thesis is structured in three articles which develop microeconometric analysis on the Italian banking industry. The first two articles investigate, relying on different theoretical frameworks and techniques, the degree of competition in the retail segment whereas the third article examines the effects of relationship banking on firm innovativeness.

The first article is entitled "Intensity of Competition and Market Structure in the Italian Banking Industry". The aim of this work is to empirically test Sutton's predictions for industry with exogenous sunk costs in the Italian banking industry. In particular, I focus my attention only on the *retail segment* since products are rather standardized and there is a limited scope for cost-decreasing or quality increasing investments.

The second article is entitled "Unit Roots and the Dynamic of Market Shares". In this article I rely on panel unit roots tests in order to infer the degree of competition in the industry. The main idea is to verify if market shares of the first five banks contain unit roots. If so, it is possible to infer that there is the chance for competitors to displace permanently the leader. On the contrary, if share are mean reverting, it is reasonable to infer that the industry is rather stable, and actors reached positions difficult to overcome.

The third article is entitled "Relationship lending and Firm innovativeness: New Empirical Evidence". The aim of this study is to test the effect of firm-bank ties on the degree of firm innovativeness using data on a sample of Italian manufacturing firms. In particular, this article departs from previous microeconometric works since it distinguishes the discovery phase from the introduction phase of new technologies.

Chapter 1

Intensity of competition and Market structure in the Italian Banking Industry

Abstract

This work tests the predictions of Sutton's model of independent submarkets for the Italian *retail banking* industry. In the first part of this paper, I develop a model of endogenous mergers to evidence the relationship between firms' conduct, market entry and market structure. In the second part, I identify the submarket dimension and estimate the relationship between market size and market structure using data on bank branches. The size of the submarkets turned out to be at most provincial whereas the limiting concentration index - as argued by Sutton for industries with exogenous sunk costs - goes to zero as the market becomes larger.

Keywords: Concentration, Truncated Poisson and Negative Binomial models, quantile regressions
JEL Classification: C24, D43, L11, L89

1.1 Introduction

Sutton's model of independent submarkets emphasizes the strategic choice of sunk costs and, unlike the traditional structure-conduct-performance approach, considers how changes in firms' conduct affects the condition of entry, altering by consequence market structure. Under this scheme, both homogeneous-horizontally differentiated products and advertising-R&D intensive (vertically differentiated products) industries can be analysed. For the former type of industries, with fixed sunk costs, it is possible to show an inverse relationship between market size and market structure. For the latter type of industries, where sunk costs are endogenous, such a negative relationship does not necessarily emerge as market size increases. This is because sunk costs, such as advertising or R&D expenditure, raise with market size. Such expenditures are choice variables of (perceived) quality: by increasing the level of advertising-R&D, firms are able to gain (or to maintain) market share. Therefore, as market size becomes larger, an '*escalation mechanism*' could raise fixed costs per firm to such an extent that the negative relationship between market size and market structure will break down. Sutton's model offers therefore very clear and testable predictions about the relationship between market-size and market-concentration. The aim to this paper is to analyse the Italian *retail banking* industry as a special case of the first type of industries, since products are rather standardized and there is a limited scope for cost-decreasing or quality increasing investments. This industry can be viewed as made of a large number of local markets that arise because there are many different geographical locations throughout the country: in every submarket products are fairly good substitutes and banks compete against each other by means of their branch locations. The degree of substitutability is substantially lower for products and services offered in neighbour submarkets. In the first part of this paper, relying on the model developed by Vasconcelos (2006), I examine the firm strategic behaviour referring to a three-stage non cooperative game. In line with Sutton's theory, the aim is to highlight the relationship between firm conduct, entry and market structure while explicitly allowing for a merger process in the industry. In so doing, it is possible to show that the incentives to merge to a monopoly are lead by the intensity of competition and by the degree of product substitution. This ultimately shows how the number of banks as well as the share of the main bank are determined in each submarket, and offers indications on the variables to be used in the empirical section. In the second part, I will estimate the market structure-market size relationship in the Italian retail banking industry. Testing this relationship empirically requires identifying a set of independent submarkets. In order to do that, I estimate the number of firms in each province using data on the national bank branch location. To take into account that the number of firms is discrete and greater than zero, a truncated Poisson and Negative Binomial models are used. This analysis confirms that the province (at most) is the size of each submarket. In fact, firm variables related to neighbour provinces turned out to be insignificant in determining the number of bank in each province. Once the size of the submarket has been identified, I will investigate the market structure-market size relationship by regressing the one firm concentration ratio on the market size variables. As the limiting concentration ratio approaches zero as market size goes to infinity, the hypothesis of exogenous sunk costs for the *retail banking* industry can be accepted. The paper is organized as follows. The next section presents the theoretical framework to analyse the relationship between firm conduct and concentration based on the Sutton approach. Sections 3 and 4 describe respectively the banking industry referring to this framework and the characteristic and the construction of the dataset. In section 5 the econometric model and results are presented. Conclusions are in the final section.

1.2 The theoretical approach

Sutton (1991, 1997, 1998) describes the impact of firm conduct on market structure identifying two key aspects: the intensity of competition and the level of endogenous sunk costs. Considering these elements, he distinguishes between two general types of industry. One class is characterized by industries that produce homogeneous and horizontally differentiated products. The other category is composed of industries engaged in the production of vertically differentiated products. In the first type of industries, the only important sunk costs are the exogenously determined setup costs, given by the technology. In such industries Sutton (1998) predicts a lower bound to concentration, which goes to zero as the market size increases and rises with the intensity of price competition. The idea is that as market size increases, profits also increase, and given free entry, other firms will enter the market until the last entrant just covers the exogenous cost for entry. Also, the higher the competition, the higher the concentration index. In fact, as the competition gets stronger, the entry becomes less profitable and the higher the level of concentration is to be in order to allow firms to cover their entry cost[1]. It is important to underline that the intensity of competition will not simply represent firm strategies but, rather, the functional relationship between market structure, prices and profits. It is derived by institutional factors, and therefore is not only captured by the price-cost margin. More generally, an increase in the intensity of competition could be represented by any exogenous influence that makes entry less profitable, e.g the introduction of a competition law (Symeodonis (2000), Symeodonis (2002)). In the second type of industries sunk costs are endogenous. Firms pay some sunk cost to enter but can make further investments to enhance their demand. As market size increases, the incentive to gain market share through advertising and R&D expenditure also increases, leading to higher fixed cost per firm. Even though room for other firms is potentially created, the 'escalation mechanism' will raise the endogenous fixed costs, possibly breaking down the negative structure-size relationship that exists in the other type of industries. For such industries Sutton's model predicts that the minimum equilibrium value of seller concentration remains positive as the market grows[2]. Sutton's model offers very clear predictions for the first group of

[1] A way to model an increase in the 'toughness of price competition' is to consider a movement from monopoly model to Cournot and Bertrand model. For any given market size, the higher the competition at final stage, the lower the number of firms entering at stage 1, and the higher the concentration index (ex-post). See Sutton (2004).

[2] To be more precise, Sutton goes further in distinguishing within the endogenous cost categories between low-α and high-α industries. In the low-α type industries, due to R&D trajectories, we will still observe low level of concentration

industries whereas it is not possible for industries where sunk costs are endogenous. Despite Sutton's insights on the relationship between market concentration and market size, there are few empirical works testing these predictions. Previous works that test the Sutton approach are Sutton (1998) for the US Cement Industry, Buzzacchi and Valletti (2005) for the Italian Motor Insurance Industry, Asplund and Sandin (1999) for the Swedish Driving Schools Sector, (Walsh and Whelan (2002)) in Carbonated Soft Drinks in the Irish retail market, Hutchinson et al. (2006) across manufacturing industries in UK and Belgium, and Ellickson (2007) for the supermarket industry in the United States. These papers mainly test Sutton's predictions by typically looking if a measure of firm inequality - such as the Gini coefficient - increases with the number of submarkets. Specific to the banking sector are the works of Dick (2007) for the Banking Industry in the United States and de Juan (2003) for the Spanish Retail Banking Sector. Although Dick (2007) investigated the relationship between market size and market concentration, she considered the banking industry without distinguishing the retail segment from the wholesale. In particular, she focused on banking quality through a set of variables, such as geographic diversification, employees compensation and branch density, finding a non-decreasing concentration ratio as market size gets larger. She concluded that endogenous quality model characterized the industry. de Juan (2003) analysed instead another important insight of Sutton's analysis: the degree of concentration and the level of aggregation of submarkets. Focusing on retail market only, and after having identified the individual submarkets, she tested the bound on the inequality of firm size distribution at different levels, local, regional and national. The purpose of this paper is to verify if empirical evidence for the Italian retail banking industry is consistent with Sutton's predictions. To apply this framework to the Italian banking industry is of interest since during the nineties it experienced a deregulation and consolidation process. Therefore, in order to identify the relationship proposed by Sutton, the choice of the year is crucial. I will assume that the industry reached in 2005, the year of the analysis, an equilibrium. Similar to de Juan (2003), I will make an effort to empirically test the size of the submarket but I will depart from her work by investigating the market size- market concentration relationship. As the focus is on the retail banking, this paper also differentiates from Dick (2007) as she considered both the industry segments, wholesale and retail. The present paper is also strictly related to the theoretical analysis developed by Cerasi (1996). Cerasi developed a model of retail banking competition in which banks compete first in branching and then in prices. In line with Sutton's analysis her model predicts that deregulation should lead to an increase in the degree of concentration whereas, with respect to branching, an increase in market size is followed by a decrease in the degree of concentration in branching.

1.3 Exogenous sunk cost industries: the model

Using the model developed by Vasconcelos (2006), as modified in order to explicitly account for the intensity of competition, this section analyses the market size-concentration relationship in exogenous sunk cost industries. In such industries, firms will face some sunk cost to enter but cannot make further investment in order to enhance their demand. Assuming that all consumers have the same utility function over n substitute goods (or n varieties of the same product) as follows:

$$U(x_1, ..., x_n; M) = \sum_k (x_k - x_k^2) - 2\sigma \sum_k \sum_{l<k} x_k x_l + M, \tag{1.1}$$

where x_k is the quantity of good k and M denotes expenditure on outside goods whose price is fixed exogenously at unity. The parameter σ, $0 \leq \sigma \leq 1$, measures the degree of substitution between goods[3]. When $\sigma = 0$ the cross product term in the utility function vanishes so that product varieties are independent in demand, whereas if $\sigma = 1$, the goods are perfect substitutes. For the utility function (1.1), the individual demand for good k is:

[3]This is a quadratic utility function and it has previously used by Spence(1976), Shaked and Sutton (1990), Sutton (1997, 1998) and Symeodonis (2000). The banking sector is usually analysed under hotelling-type model. However, it is possible to show that any hotelling-type model is a special case of vertical production differentiation. See Cremer and Thisse (1991).

$$p_k = 1 - 2x_k - 2\sigma \sum_{l \neq k} x_l \tag{1.2}$$

If there are S identical consumers in the market and we denote with x_k the per-capita quantity demanded of good k, market demand for this good is Sx_k.

Considering now a three stage game. In the first stage, a sufficiently large number of ex-ante identical firms, N_0, simultaneously decide whether or not to enter the market incurring an entry cost of ϵ. In the second stage, firms that have decided to enter decide to join a coalition. All the firms that have decided to join the same coalition then merge. In the third stage, firms set their output. All coalitions are assumed to face the same marginal cost of production c, which we can normalize to zero.

1.4 The game: equilibrium analysis

In stage 2, each firm $i \in \{1, ..., N\}$ simultaneously announces a list of players that it wishes to form a coalition with. Firms that make exactly the same announcencement form a coalition together. For example, if firms 1 an 2 both announced coalition $\{1, 2, 3\}$, while firm 3 announced something different, then only players 1 and 2 form a coalition. Since all firms are initially symmetric, members of each coalition are assumed to equally share the final stage profit.

Let $\lambda = \frac{dx_j}{dx_i}$ represent firm i's conjectural variation, that is its expectation about the change in its competitors production resulting from a change in its own production level, and assume that this conjecture is identical for all firms ($\lambda_i = \frac{dx_j}{dx_i} = \lambda$).

Assuming that quantity is a strategic variable, profit maximization requires that $\partial \Pi_i / \partial x_i = 0$. In equilibrium:

$$x_i = \frac{1}{2(2 + (N-1)\sigma(1+\lambda))} \tag{1.3}$$

and the profit of each of the N firms is

$$S\Pi_i = S\frac{1+\lambda(N-1)\sigma}{2(2+(N-1)\sigma(1+\lambda))^2} - F \tag{1.4}$$

Let Λ be equal to $\sigma(N-1)\lambda$. It is possible to refer to Λ as the competitive intensity of the industry, with lower values of Λ corresponding to more intense competition.

For $F \geq 0$, $N \geq 2$ and $-1 \leq \Lambda \leq 1$, and $0 \leq \sigma \leq 1$ each firm's profit is a decreasing function of the number of firms in the industry, its competitive intensity and the amount of fixed costs. Two reasons could lead firms to merge: market power and efficiency. To maintain things simpler, I avoid to account for efficiency gains. In this analysis, firms could not make any further investments to enhance their quality (and hence the demand) of the product offered. So, it possible to set $F = 0$. In any case, a clear picture in similar framework is offered by Rodrigues (2001).

Following the traditional backward induction procedure, I analyze the condition under which I get a monopoly in *exogenous sunk cost industries* model.

Quantity setting stage Let N_2, $N_2 \leq N \leq N_0$, denote the number of coalitions of firms at the end of stage 2. From equation (1.4) firm profits are

$$S\Pi(N_2) = \frac{1+\Lambda}{2(2+\sigma(N_2-1)+\Lambda))^2} \tag{1.5}$$

Coalition formation stage At this stage those firms who entered may merge to form a coalition. A coalition structure is said to be an outcome of a Nash equilibrium if no player has incentive to either (individually) migrate to another coalition or to stay alone (Vasconcelos (2006); Yi (1997))[4]. Consider a coalition structure composed of coalitions of the same size. It is said to be stand-alone stable if

$$\frac{N_2}{N}[S\Pi(N_2|\Lambda,\sigma)] > S[\Pi(N_2+1)|\Lambda,\sigma] \tag{1.6}$$

[4]To be more precise, this latter case in which no firm can unilaterally improve its payoff by forming a singleton coalition is called stand-alone stability. However, stand-alone stability is a necessary condition for Nash stability.

In case of monopoly, $N_2 = 1$. Hence, in order for a single 'grand coalition' to be the outcome of a Nash equilibrium of the coalition formation game in *exogenous sunk cost industries*, the following is a necessary and sufficient condition[5]

$$\Pi(1)/N > \Pi(2) \tag{1.7}$$

Hence,

$$\frac{(1+\Lambda)}{2(2+\sigma+\Lambda))^2} < \frac{1}{8N}$$

$$N < \frac{(2+\sigma+\Lambda))^2}{4(1+\Lambda)} \equiv \bar{N}(\sigma,\Lambda) \tag{1.8}$$

I restrict the industry conjectural variation coefficient to the range $-1 \leq \Lambda \leq 1$. In so doing, the possibility of Λ being larger than the value that would imply perfectly collusive post-merger behaviour is restricted.

A merger towards monopoly leads to the formation of a single grand coalition with N firms. A firm belonging to the initial wave of N entrants will get a share $1/N$ of the coalition overall profits, whereas by free-riding on its N-1 merging rivals it can obtain duopoly profits. Each time in which the 'grand coalition' is unstable, as market size increases, more firms want to enter and to free ride and form a duopoly instead of joining the grand coalition. That means, as the market size rises, the concentration ratio goes down [6]. This result shows how this process in turn can affect the one firm concentration ratio, $C_1 = \frac{q}{N_2 q} = 1/N_2$.

When Λ lies in the range previously defined, and fixed costs are zero, $N(\sigma,\Lambda)$ is strictly decreasing in Λ. Therefore, the weaker the competitive intensity, the larger the pre-merger market concentration should be for a monopoly to emerge through merger.

In particular, if $\sigma = 1$, we can rewrite equation (1.8) as

$$N < \frac{1(\Lambda+3)^2}{4(\Lambda+1)} \tag{1.9}$$

[5]The only possible deviation it is in fact towards the singleton coalition.

[6]It is valuable to remark that in this model it is implicitly assumed that the pre-merger behaviour is not affected by the coalition formation stage.

The RHS is strictly decreasing in Λ. As Λ approaches -1, the value of perfect competition, condition (1.9) is alway satisfied, and so, merger to monopoly would occur whatever the number of firms in the industry. Hence, the higher the intensity of competition at stage 3, the lower the pre-merger market concentration could be in order for a monopoly to emerge through merger. When $\Lambda = 0$, that is firms behave as in Cournot, monopolization will occur only if $\sigma \geq 0.83$. If this condition is not met and more than two firms enter in stage 1, and merge in a single grand coalition, that equilibrium might not be stable. As σ approaches 1, competition becomes tougher as products are closer substitutes, and the lower bound to the one firm concentration ratio decreases as market size increases. On the other hand, in perfectly cooperative industries, where $\Lambda = 1$, or when demands are perfectly independent, where $\sigma = 0$, merger to monopolization will never occur. However, it is important to remark that we are not considering cost efficiency gains that would probably give an incentive to merge even in the case that market demands are completely independent.

Entry stage At stage 1 firms decide to enter.

If $\sigma = 1$ products are perfect substitutes, a merger to monopoly will occur at the second stage of the game if firms compete very toughly. Then, if firms anticipate that a monopoly coalition structure is going to be formed at stage 2, firms will enter up to a point at which N is the largest integer value satisfying

$$\frac{1}{N}(S\Pi(1)) \geq \epsilon \tag{1.10}$$

where $\epsilon > 0$ is the entry fee. By the same reasoning, therefore, if the competitive intensity is extremely strong, the firms will merge to monopoly. For any given level of market size, the equilibrium level of concentration is higher. However, entry will occur at the first stage and the lower bound to concentration goes down[7]. If products are imperfect substitutes - and $\Lambda = 0$ - a merger to monopoly might not take place. In particular, when

[7] Also, from the previous analysis, since $\partial\Pi/\partial N < 0$ and $\partial\Pi/\partial\Lambda > 0$, by applying the implicit function theorem, one concludes that $\partial N/\partial\Lambda = \frac{-\partial\Pi/\partial\Lambda}{\partial\Pi/\partial N} > 0$. The equilibrium number of firms is decreasing in the intensity of competition at stage 3.

$\sigma < 0.83$, a merger to monopoly might not take place since a firm could prefer to get all the profits of a duopolist. This means that as the market size rises, more firms enter and this makes the monopoly unsustainable as individual firms want to free ride and form a duopoly. Thus, there is an upper bound to concentration that goes down as market size increases (Vasconcelos (2006)).

1.5 The Italian retail banking industry

The presence of different territorial dynamics is a characteristic of the Italian banking industry (Guiso et al. (2004, 2006); Colombo and Turati (2004)). I consider the retail Italian banking industry as belonging to an industry of the first type, where sunk costs are exogenous and the size of the submarkets is provincial. Since lending and borrowing take place mostly in a narrow geographical place and operation are similar and repeated during time, this industry can in fact be viewed as made of a large number of local markets, corresponding to different provinces (geographical units close to US counts). These submarkets are independent both from the supply and demand side. On the supply side, in each one of these independent submarkets, banks' goods are fairly substitute whereas banks' products of neighbouring provinces are not. In particular, in each province banks can mitigate price (interest rate) competition by means of their branch location[8]. However, opening new branches, independently of the size of their operations, has fixed costs, for example the cost of hiring personnel, the cost of renting or buying facilities in particular province and other province specific elements. As documented in (Cerasi et al. (2000)), in Italy in the recent years, as a result of reforms on entry and branching regulation, the cost of branching has decreased. On the demand side, despite the advances in home and phone banking, preferences of customers seem to be still biased toward entities with strong regional and local contents. A customer is likely to shop only at those banks that operate in the local area where he lives and works. In other words, zero/small cross-elasticities are likely to characterize the demand of geographically separated submarkets

[8] See also Cerasi et al. (2002) and Cohen and Mazzeo (2004)

whereas positive elasticities are likely to characterize the demand in each province.

1.6 Exogenous or endogenous sunk costs?

As we would expect both exogenous and endogenous sunk costs to be relevant in the banking industry, with both horizontal and vertical differentiation, some point of remarks are deserved. In this work I am considering the *retail sector* by looking at branches as the main distributional channel of certain standardized banking products. Therefore, I am not looking as in Dick (2007) at branches, as one of the costs in advertising and quality (employee compensation, branch staffing) that banks will incur in order to enhance consumer willingness to pay. Indeed, as banks become more and more visible through branches, one could consider branches as a form of advertising itself. In other words, I am assuming that branches of different banks offer similar (bundle of) services despite bank size and, hence, the number of branches in a given submarket could be considered as the number of varieties of services offered by banks. Even in the case, however, there are circumstances in which endogenous costs could arise. As pointed out by Petersen and Rajan (1995) relationship lending may generate severe barriers to entry. However, the advent of information and communication technologies increased the ability of banks to open branches in distant locations, considerably reducing the cost of distance-related trade and enhancing competition in local banking markets (Berger and Udell (2006), Affinito and Piazza (2005))[9]. In addition, developments in the financial industries with new contracts and new intermediaries are likely to reduce the role of close bank-firm relationships (Rajan and Zingales (2003)). The opinions are not unique. Whatever the conclusion might be, we can foresee that it will at least influence the structure of the banking system in terms

[9] Besides, Berger himself has recently taken an opposite view with respect to his previous study (Berger et al. (2003)) where it is claimed that services to small firms are likely to be provided by small banking institutions since they meet the demands of informationally opaque SMEs that may be constrained in the financing by large institutions. He now claims that this vision could be an oversimplification: new transaction technologies are now available enabling large banks to overcome informational constraints.

of the local nature of the banks but not the number of branches that could be opened given market demand[10]. It is obvious that in the industry as a whole (*retail* and *wholesale*) both endogenous and exogenous interact costs with one another to determine market structure. The approach and conclusion could be very different (Dick (2007)).

1.7 Market equilibrium

The predictions of Sutton's model apply to markets in equilibrium. However, discontinuities in the normative (or economic) conditions can lead to process of consolidation. Unless when are observing the market at the end of this process, it will be difficult to disentangle the relationship between competition and concentration as predicted by Sutton from that caused by mergers and acquisitions. This means that I am making the implicit assumption that the retail baking sector reached an equilibrium in 2005, the year for which I collected the observations. This assumption - though strong - seems reasonable. Beginning in the 1980s, the Italian Banking system underwent a series of reforms aimed at increasing the competition in the market through liberalizing branching and easing the geographical restrictions on lending. In fact, the opening of new branches had been regulated by the 'branch distribution plan', issued every four years. The last distribution plan was issued in 1986 and, since March 1990, the establishment of new branches has been completely liberalized. The number of branches increased steadily, up to 31.081 in 2005, as well as the number of people served by each branch, 47 per 100.000 inhabitants in 2004 (compared to 59 EU mean). In particular, the number of banks mergers and acquisitions of control per year was 45 in 1990 and decreased substantially to 5 in 2005[11]. At the same time, in more than 50% of the provinces, new banks entered the market. This process of new entry, parallel to the pro-

[10] To have a picture of the role of local banks and how the probability of branching in a new market depends on the features of both the local market and the potential entrant, see Di Salvo et al. (2004), Bofondi and Gobbi (2004) and Felici and Pagnini (2005).

[11] Referring to December 2005. It is important to remark that the process of consolidation with foreign banks is now gaining relevance. See ICB (2004).

cess of consolidation, made the average number of banks in each province rise from 29 in 1990 to 34 in 2005.

1.8 Characteristics and construction of the dataset

The dataset is composed of 103 Italian provinces and 784 banks. In total, there are 85 groups of banks to which 230 banks belong. The greater part of banks, 554, does not belong to any group. The Italian territory is divided into 20 regions and 103 provinces, which are geographical units close to US counties. For each provinces, I have data on the number of banks and their number of branches for the year 2005 as collected by the Italian Central Bank (Banca d'Italia)[12]. I also have data about GDP, number of inhabitants, density of population as collected by National Institute of Statistics (Istat). According to the criteria developed below, four provinces will be excluded when estimating the submarket size since these are - by definition - considered 'isolated' provinces[13]. A description of the variables involved in the analysis follows, as well as indications for the theoreticals variables they should account for. The name of the variable that will be used in the empirical assessment is reported in square brackets. Summary statistics are reported in tables (1).

- *Concentration* $= C_1$

  To measure concentration the 'one-bank concentration ratio', $[C1]$, is used. The bank concentration ratio is defined as the fraction of the number of branches owned by the largest bank within the market.

- *Market size* $= S$

[12] `http://siotec.bancaditalia.it/sportelli/main.do?function=language\&language=ita`.

[13] These provinces are: Potenza, Palermo, Trapani and Sassari. Therefore, in that case I considered 573 banks or group of banks over 99 provinces for a total of 2673 observations. I do not consider in this count the number of branches belonging to foreign banks. For further information, see ICB (2005) and `http://www.bancaditalia.it/pubblicazioni/ricec/relann/rel05/rel05it/vigilanza/rel05_attivita_vigilanza.pdf`

It is likely to vary with the level of demand measured by GDP, [VA_pct], and by population, [$logPOP$], in the province considered[14].

- *Intensity of competition and product differentiation* $= \Lambda$ *and* σ

  So as to control for different market features, I control for population density, [$DENS$], measuring thousands of people per Km^2. The higher the density, the lower the number of banks: comparing two submarkets with the same number of inhabitants, I expect that the number of branches will be less in the submarket with a high population density.

  To measure the intensity of competition and product differentiation, I computed three indices:

  - [K] $= Totalbranches/Km^2$. It represents the monopolistic power of each branch and could be considered as a proxy of the (inverse of) transportation costs. More branches in the same provinces means, for each consumer, a lower distance to cover to reach a branch, a weaker power exerted by bank branch and an overall higher degree of competition.

  - [P] $= Totalbranches/Population$. It is the number of branches for a thousand inhabitants. The higher P, the higher the competition. It can be considered as a proxy for the (inverse of) queueing costs. The less the population served by each branch (or the higher the number of branches for each individual), the lower the cost met by the customers[15].

  - [CV] $=standarddeviation/Branchesmean$. It is the coefficient of variation. It is a dimensionless number and it is calculated by dividing the standard deviation by the mean of branches in each

[14] Since data on GDP for the year 2005 was not available,in the analysis I used the percentage of value added pertaining to each province for year 2004. The relative position of each province is unlikely to markedly change from one year to another. Regarding data on population for the year 2005 I relied on Istat forecasting at `http://demo.istat.it/stimarapida/`

[15] It is interesting to note that these two indices, K and P, split the information contained in the density of population, $DENS = population/Km^2$

province. The higher the CV, the higher the degree of differentiation by branches opening, since some bank has smaller branch network size whereas others have greater branch network size.

- *Market Borders*

  Since the unit of observation is the bank (or group of banks), I also compute for each bank in every submarket (province)

  - the total number of its own branches [$NB\_OWN_{i_m}$]

  - the total number of branches of its competitors [$NB\_COMP_{i_m}$]

  The same quantities are also computed for all the 'closest' provinces (less than 100 Km) [$NB\_OWN\_OUT_{i_m}$] and [$NB\_COMP\_OUT_{i_m}$][16].

1.9 Intensity of competition and concentration: Empirical model and results

According to Sutton's model, the number of branches per submarket is a function of the relative size of the submarket, of the intensity of competition and of the cost incurred to entry. As market size increases, profits also increase, and given free entry, other firms will enter the market until the last entrant just covers the exogenous cost of entry. As the previous analysis also showed, the relationship between the number of firms (or concentration) and the market size will in general depend on the intensity of competition and the degree of product differentiation.

The Italian Antitrust Authority defines the province as the relevant market. Prior to analyse the relationship between the one-bank firm concentration ratio and market size, this hypothesis is tested.

[16] I performed an alternative analysis computing the number of branches of each bank, and those of its competitors, outside the province but in the same region. The reason for trying this specification is to test the alternative regional dimension for market size that is, in general, used by the authorities or in similar studies. The results are substantially analogous.

1.10 Model description: identifying the size of the submarket

In order to test the submarket dimension, I construct a model for the number of competitors in each province. Given the data, no observations are possible for provinces with zero banks, since a criterion for sample inclusion is that there is at least one banks in the province. This is to be distinguished from datasets without 0 values, but which may have 0s. Thus, the dependent variable of the model, the number of banks in each province, is truncated at zero, taking only positive values. A zero-truncated Poisson and Negative Binomial models are therefore appropriate, since these models allows us to take into account that the dependent variable, $NFIRMS$, is also a non negative-integer. The truncated densities of these models are easily obtainable by slightly modifying the untruncated models and have been presented in Cameron and Trivedi (1998), Gurmu and Trivedi (1992) and Gurmu (1991).

The latent variable, $NFIRMS^*$, is assumed to be

$$NFIRMS^*_{i_m} = X'_{i_m}\delta + e_{i_m} \tag{1.11}$$

where $m = 1...99$ is the submarket, i_m is bank i in submarket m,

$X_{i_m} \equiv (NB\_OWN_{i_m}, NB\_COMP_{i_m}, NB\_OWN\_OUT_{i_m},$
$NB\_COMP\_OUT_{i_m}, CV_m, P_m, K_m, VA_m\_pct)$

and

$$NFIRMS_{i_m} = NFIRMS^*_{i_m} \quad if \quad NFIRMS^*_{i_m} > 0 \tag{1.12}$$

Since not all of the 573 banks (or group of banks) are active in every province, the subscript i_m goes, for each m, from $N_{m-1}+1$ to N_m, where the total numbers of banks, $N_m = N_{m-1} + n_m$, gets incremented by n_m, the total number of banks in each province and $N_0 = 0$. The overall sample size, $n_1 + ... + n_{99}$, is equal to 2673. Observations may be considered independent across provinces (clusters), but not necessarily within groups. Cluster devices must be adopted. The number of bank branches in each province, $NB\_OWN_{i_m}$ is likely to vary with the level of demand. Therefore, it is reasonable in the estimation to control for the level of demand, represented by GDP, $[VA_m\_pct]$, and the population spread, $[DENS_m]$. Furthermore, to account for different intensity of competition in the province, I computed two indices of competition, K_m and P_m[17]. Then, to take into account the border of submarkets, I consider the number of branches of each bank in each province $[NB\_OWN_{i_m}]$ and outside the province $[NB\_OWN\_OUT_{i_m}]$, and the number of branches of 'other banks', distinguishing them between competitors' bank branches in the same provinces $[NB\_COMP_{i_m}]$ and competitors' bank branches outside the provinces $[NB\_COMP\_OUT_{i_m}]$[18]. The degree of product differentiation is captured by the coefficient of variation, $[CV_m]$, that measures how banks are differentiated in terms of total size of their network of branches inside each province. For the Poisson model the probability that there are exactly N firms in the market, conditional on N being greater than zero, is

$$Prob(NFIRMS_{i_m} = N|N > 0) = \frac{e^{-\gamma_{i_m}}\gamma_{i_m}^N}{N!(1 - e^{\gamma_{i_m}})}, \tag{1.13}$$

for $N = 1, ..., \infty$ and $\gamma_{i_m} = exp(\delta X_{i_m})$.

Unlike the Poisson distribution, the zero-truncated Poisson distribution does not present equidispersion (that is, the equality between the conditional mean and variance). In fact, the average of the truncated distribution is higher than the average of non-truncated distribution while its variance is smaller. In addition, contrary to the non-truncated case (Asplund and Sandin (1999)), the estimates of the regression parameters will be biased and inconsistent in the presence of overdispersion because consistency requires the proper specification of all the moments of the underlying relevant cumulative distribution. These findings are similar to the result that the Tobit estimator, unlike ordinary least squares, yields inconsistent parameter estimates in the presence of heteroscedasticity (See Grogger and Carson (1992)). Given the importance of accounting for overdispersion in

[17]See section 4.

[18]Please see note 14.

the truncated context, I also present a model for truncated counts based on the Negative Binomial distribution. The conditional distribution of a truncated Negative Binomial is

$$Prob(NFIRMS_{i_m} = N|N > 0) = \frac{\Gamma(N + \frac{1}{\alpha})}{\Gamma(N+1)\Gamma\frac{1}{\alpha}} * \frac{1}{(1+\alpha\gamma)^{\frac{1}{\alpha}} - 1} * (\frac{\alpha\gamma}{1+\alpha\gamma})^N, \tag{1.14}$$

As for the Poisson distribution, the average of the truncated negative binomial distribution is higher than that of the non-truncated one. Though the truncated Poisson distribution no longer shows the equidispersion characteristic, the truncated Negative Poisson model introduce overdispersion, in the sense that its variance is higher than that of the Poisson[19].

1.11 Model description: testing market size-market concentration relationship

In practice the relationship between market size and market concentration has been investigated by estimating a lower bound where a concentration measure is regressed on market size variables. In that case,

$$log(\frac{C_1}{1-C_1}) = a + b\frac{1}{log(S/\epsilon)} + v \tag{1.15}$$

The constant represents the value of the limiting concentration as the market size approaches infinite, that is $C_\infty = exp(a)/1 + exp(a)$. The most used approach is the Smith's two step procedure, where the error distribution is a two or three parameters assumed to be drawn from a Weibull distribution. See for example Marìn and Siotis (2007). Lyons and

[19]For the truncated Poisson distribution the first and the second moment are:

- $E(N|N > 0; X) = u = \lambda + \sigma$
- $V(N|N > 0; X) = \sigma^2 = \lambda - \sigma(u - 1)$

with $\sigma = \lambda/(e^\lambda - 1)$. The mean and the variance of the truncated negative binomial regression are the following:

- $E(N|N > 0; X) = u^* = \lambda + \sigma^*$
- $V(N|N > 0; X) = \sigma^*2 = \lambda + \alpha\lambda - \sigma^*(u^* - 1)$

with $\sigma^* = \lambda/((1 + \alpha)^{\alpha^{-1}\lambda} - 1)$.

Matraves (1996) proposed to use a stochastic (cost) frontier approach, allowing for disequilibrium deviations from the bound. As estimating a stochastic lower bound by maximum likelihood methods is possible only when the least squares residuals are positively skewed, Symeodonis (2000) suggest to simply use OLS regressions. I will follow Giorgetti (2003) relying on quantile regression, as estimations obtained with this procedure are robust to outliers. In particular, I will estimate the following lower bound

$$\begin{aligned} log(\frac{C_1}{1-C_1}) = \quad & a + a_1 TYPE1 + a_2 TYPE2 + a_3 REGION1 \\ & + a_4 REGION2 + b\tfrac{1}{log(S/\epsilon)} + v \end{aligned} \tag{1.16}$$

where TYPE and REGION are dummies variables, which account for different macro regions in which the Italian territory can be divided and for the different type of banks. For example, the limiting concentration ratio in a province in the NORTH, where the major bank is a cooperative, will be equal to $C_\infty = exp(a + a_1 + a_3)/1 + exp(a + a_1 + a_3)$.

1.12 Results

The results of the zero truncated Poisson are reported in table (2). These results suggest that province could be considered - in general - as an independent submarket. As expected, the value of the coefficient is higher for branches in the same provinces [$NB\_OWN_{i_m}$] and [$NB\_COMP_{i_m}$], and lower and close to zero for banks outside [$NB\_COMP\_OUT_{i_m}$]. Regression in column 2 and 3 in table (2) replicate the analysis in column 1 accounting for *i)* different macro-regions ($REGION1$=Nord,$REGION2$=Centre, $REGION3$=South) in which it is possible to group provinces, and *ii)* different types of banks ($TYPE1$=BCC, $TYPE2$=BP, $TYPE3$= S.p.A). The inclusions of these variables improve the explanatory power of the model (likelihood ratio tests are significant). In particular, supporting the point of independence among provincial submarkets, we can accept the null hypothesis that both the coefficients of branches belonging to banks outside the province are zero. The sign for the K coefficient is negative and significant whereas the value of the P coefficient is smaller, positive and significant. These results suggest, as one could expect, that transportation costs are more relevant in the retail market, and, therefore, a higher branch density increases competition lowering the expected (ex-post) number of banks. The value of the coefficient on CV is positive and significant. In accordance with the model developed in the previous section, the higher the degree of differentiation, the higher the number of banks. Since consumers have preferences about total number of branches, some banks have greater network size with respect to other competitors and are able to capture more consumers by differentiating themselves by opening more branches. In equilibrium, therefore, higher asymmetry in branch size (a higher value of CV) is compatible with a large number of banks. On the contrary, the sign of the coefficient for the density of population is significant with unexpected signs, where GPD is positive and insignificant. This is probably due to the non-linear relationship between these variables and the dependent variable, and to the fact that higher density will capture the same effect of GDP (since higher density is associated with higher GDP). Results for the zero Truncated Negative Binomial are reported in table (3). These regressions are in line with those of the zero Truncated Poisson. However, our interest lies in measuring the change in the conditional mean of $NBANKS$ when regressors X change by one unit, the so called marginal effects[20]. For reporting purposes, in tables (4) a single response value - the mean of the independent variables - is used to evaluate the marginal effects for regression 3 in table (2) and (3). At this point it is important to control for overdispersion, since in context with truncation and censoring it leads to problems of inconsistency (Hilbe (2007)). Several test procedures can be followed to test the (truncated and untruncated) Poisson model against the Negative Binomial model. As the Negative Binomial model degenerate into a Poisson model when $\alpha = 0$, all tests (score test, Wald test, likelihood test) are based on testing the overdispersion parameter α equal to zero (Yen and Adamowicz (1993)). Since both the Poisson and the Negative Binomial model have been estimated a likelihood-ratio test

[20]For linear regression marginal effects coincide with the estimated coefficients. For non linear regression this is no longer true. In that case, $E[NBANKS|X] = exp(X'\beta)$, then $\partial E[NBANKS|X]/\partial X = exp(X'\beta)\beta$ is a function of both estimated parameters and regressors.

is straightforward. From the previous tables, it is possible to compute the likelihood-ratio test of $\alpha = 0$. This is the likelihood-ratio chi-square test that is equal to $\chi^2_{(1)} = -2(ll(Poisson) - ll(NegativeBinomial))$. The large test statistic would suggest that the response variable is over-dispersed and it is not sufficiently described by the Poisson distribution. In all cases, this test is significant (for example, for regressions (3) is equal to $\chi^2_{(1)} = 357.26$, $p < 0.001$). In the end, these preliminary analysis allows me to analyse the main relationship of interest, that is the one between the size of the market and the one-firm concentration ratio at the provincial level, relying only on a subsample made of one observation for each submarket. I will measure market size by means of two variables: the population in each province and the number of banks as estimated in previous section. Concerning the choice of this latter variables, it important to notice that in homogeneous industries, the number of firms represents the ratio between market size and sunk costs, that is S/ϵ . In addition, I can rely on 99 submarkets instead of 103 since four provinces has been excluded from the previous regressions as were considered isolated by definition. Tables (5) and (6) report quantile regressions for the fifth, tenth and fiftieth percentile. The results are very similar and support the hypothesis that the retail banking industry is characterized by exogenous sunk costs: the estimated limiting concentration C_∞, approaches zero as the market size approaches infinite. It would be better to have an industry with endogenous costs so as to compare the value of the limiting concentration. However, the quantile regressions, using both measures of market size, indicate that when market size increases, the concentration index goes down. This result is weaker in provinces located in the South and in the Centre whereas it is stronger in province where the main banks is a TYPE2 (that is, Banche Popolari).

1.13 Robustness checks

The aim of this section is to control for issues that could weaken previous results, mainly endogeneity and model specifications. Endogeneity may be an important concern when testing the size of each submarket, since there are variables that could be considered jointly determined with the number of banks if the industry has not reached an equilibrium. In particular, at firm level, a troubling variable could be the number of branches a bank has in the province (NB_OWN). The easiest way to test for endogeneity is to use a method suggested by Wooldridge (1997) for count models with endogenous explanatory variables along similar lines to those suggested in other limited dependent variable contexts by Smith and Blundell (1986) and Rivers and Voung (1988). For any given explanatory variable x which is potentially endogenous, it is possible to estimate a reduced form regression of the form

$$x = z'\pi + v \tag{1.17}$$

where z represents a vector of exogenous explanatory variables including at least one not included in x for identification, π the vector of reduced form coefficients and v is the reduced form error term. If it is possible to obtain consistent estimates of π, Wooldrige shows that the residuals $\hat{v} = x - z'\hat{\pi}$ can be included as an additional covariate in a maximum likelihood estimator for count data model. A significant coefficient on $\hat{v}$ in the augmented regression is a robust test of endogeneity of x. I test for possible endogeneity of NB_OWN using as an identifying instrument the same variable in year $t-2$. The reason to choose NB_OWN_{t-2} instead of in year NB_OWN_{t-1} is to avoid the risk of unit root. The reduced form is presented in table (7), whereas the residuals from this regression are then used as an additional covariate in the zero truncated Negative Binomial regression in table (8). The coefficient of the residuals is not significant, suggesting that in year 2005 the Italian Banking industry reached an equilibrium. To analogous (not reported) conclusions leads a test for endogeneity of NB_OWN and NB_COMP. Another concern is related to the specification on the model for the relationship between market size and the one firm concentration ratio. A better alternative to estimate a model where the dependent variable is a proportion is to use generalized linear models (Papke and Wooldridge (1996)). Results reported in table (9) confirm those obtained by means of standard regression model .

1.14 Conclusions

The aim of this work was to test Sutton model of independent submarkets checking his predictions for the Italian *retail* banking industry and using the framework for the exogenous sunk costs industries. Even though the banking industry as a whole should be considered as characterized by endogenous sunk costs, there are several features that indicate the retail industry to be one of the former type. In particular, as banks branches sell slightly differentiated products in the retail sector, it is possible to look at the number of banks branches as different varieties of the same product offered by banks to their client. In addition, despite the advances of the phone banking, consumers' preferences are still biased toward regional entity, suggesting province as submarket dimension.

The model developed in the first part of the paper indicates which factors should influence the number of banks in each submarket, and as a consequence the one firm concentration ratio: the initial number of banks, the intensity of competition and the degree of product differentiation.

In the second part, a truncated Poisson and Negative Binomial model have been used in order to estimate the number of banks in each submarket. This way of proceeding allowed me to check the hypothesis about the size and the independence among submarkets. In fact, the value of the coefficient on the number of branches for banks outside the provinces, but within a radius of a hundred of kilometers, turned out to be insignificant. These results permitted to examine the one bank concentration ratio at provincial level. Interestingly, the limiting concentration ratio approaches zero as market size goes to infinity. That means that exogenous sunk costs are involved in the Italian *retail* banking industry. As argued by Sutton, as the dimension of the submarket becomes larger, and given free entry, the value of concentration ratio has to go down.

References

Affinito, M. and Piazza, M. (2005), "What determines banking structure in European regions?" *Paper prepared for the XIV International Tor Vergata Conference on Banking and Finance.*

Asplund, M. and Sandin, R. (1999), "The Number of Firms and Production Capacity in Relation to Market Size," *The Journal of Industrial Economics*, 69–85.

Berger, A.N. and Udell, G. (2006), "A More Complete Conceptual Framework for SME Finance," *Journal of Banking and Finance*, vol. 30, 2945–2966.

Berger, A., Dai, Q., Ongena, S. and Smith (2003), "To What Extent Will the Banking Industry be Globalized? A study of bank nationality and reach in 20 European nations," *Journal of Banking and Finance*, vol. 27, 383–415.

Bofondi, M. and Gobbi, G. (2004), "Bad Loans and Entry into Local Credit Markets," *Banca d'Italia Temi di discussione, No 509.*

Buzzacchi, L. and Valletti, T. (2005), "Firm Size Distribution: Testing the Independent Submarkets Model in the Italian Motor Insurance Industry," *International Journal of Industrial Organization*, vol. 24, 1–50.

Cameron, A.C. and Trivedi, P.K. (1998), *Regression Analysis of Count Data*, Cambridge University Press.

Cerasi, V. (1996), "A model of Retail Banking Competition," *mimeo.*

Cerasi, V., Chizzolini, B. and Ivaldi, M. (2000), "Branching and Competitiveness across Regions in the Italian Banking Industry," in "Industria Bancaria e Concorrenza," Il Mulino, a cura di Polo, M.

Cerasi, V., Chizzolini, B. and Ivaldi, M. (2002), "Branching and Competition in the European Banking Industry," *Applied Economics*, vol. 34, 2213–2225.

Cohen, M. and Mazzeo, M. (2004), "Market Structure and Competition Among Retail Depository Institutions," *FEDS Working Paper, No 4.*

Cremer, H. and Thisse, J. (1991), "Location Models of Horizontal Differentation: a Special Case of Vertical Differentiation, Models," *The Journal of Industrial Economics*, vol. 39, 383–390.

de Juan, R. (2003), "The Independent Submarkets Model: an Application to the Spanish Retail Banking Market," *International Journal of Industrial Organization*, vol. 21, 1461–1487.

Di Salvo, R., Lopez, J.S. and Mazzilis, C. (2004), "I Processi Concorrenziali e l'Evoluzione del Relationship Banking nei Mercati Creditizi locali," in "La Competitività dell'Industria Bancaria Italiana," Nono Rapporto sul Sistema Finanziario Italiano. Fondazione Rosselli, a cura di Bracchi G. e Masciandaro D., pages 87–118.

Dick, A.A. (2007), "Market Size, Service Quality and Competition in Banking," *Journal of Money, Credit and Banking*, vol. 39, 49–81.

Ellickson, P. (2007), "Does Sutton apply to supermarkets?" *The RAND Journal of Economics*, vol. 38, 43–59.

Felici, R. and Pagnini, M. (2005), "Distance, bank heterogeneity, and entry in local banking market," *Banca d'Italia Temi di discussione, No 557.*

Giorgetti, M. (2003), "Lower Bound Estimation - Quantile Regression and Simplex Method: An Application to Italian Manufacturing Sectors," *Journal of Industrial Economics*, vol. 51, 113–120.

Grogger, J.T. and Carson, R.T. (1992), "Models for Truncated Counts," *Journal of Applied Econometrics*, vol. 6, 225–238.

Gurmu, S. (1991), "Tests for Detecting Overdispersion in the Positive Poisson Regression Model," *Journal of Business and Economic Statistics*, vol. 54, 347–370.

Gurmu, S. and Trivedi, P. (1992), "Overdispersion Test for Truncated Poisson Regression," *Journal of Econometrics*, vol. 54, 347–370.

Hilbe, J. (2007), *Negative Binomial Regression*, Cambridge University Press.

Hutchinson, J., Konings, J. and Walsh, P. (2006), "Company Size, Industry and Product Diversification. Evidence from the UK and Belgium," *mimeo*.

ICB (2004), "Relazione Annuale," *Banca d'Italia*.

ICB (2005), "Relazione Annuale," *Banca d'Italia*.

Lyons, B. and Matraves, C. (1996), "Industrial Concentration," in "Industrial Organization in the European Union: Structure, Strategy, and the Competitive Mechanism," Oxford University Press, Lyons et al. (eds).

Marìn, P. and Siotis, G. (2007), "Innovation and Market Structure: an Empirical Evaluation of the 'Bound Approach' in the Chemical Industry'," *Journal of Industrial Economics*, vol. 55.

Papke, L. and Wooldridge, J. (1996), "Econometric methods for fractional response variables with an application to 401(k) plan participation rates," *Journal of Applied Econometrics*, vol. 11, 619–632.

Petersen, M. and Rajan, R. (1995), "The Effect of Credit Market Competition on Lending Relationship," *Quarterly Journal of Economics*, vol. 110, 407–43.

Rajan, R. and Zingales, L. (2003), "Banks and Markets: The Changing Character of European Finance," *NBER Working paper series*.

Rivers, D. and Voung, Q. (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics*, vol. 39, 347–366.

Rodrigues, V. (2001), "Endogenous mergers and market structure," *International Journal of Industrial Organization*, vol. 47, 1245–1261.

Smith, R. and Blundell, R.W. (1986), "An Exogeneity Test for a Simultaneous Equation Tobit-Model with an Application to Labor Supply," *Econometrica*, vol. 54, 679–685.

Sutton, J. (1991), *Sunk Costs and Market Structure*, MIT Press.

Sutton, J. (1997), "Game Theoretic Models of Market Structure," in "Advances in Economics and Econometrics: Theory and Applications," Cambridge Univ. Press, Kreps, D. M. and Wallis, K. F (eds).

Sutton, J. (1998), *Sunk Costs & Market Structure: Theory and History*, MIT Press.

Sutton, J. (2002), "Market Structure. The Bound Approach," LSE, Paper prepared for the Volume 3 of the Handbook of I.O.

Symeodonis, G. (2000), "Price Competition and Market Structure: The impact of Cartel Policy on Concentration in the UK," *The Journal of Industrial Economics*, vol. 48, 1–26.

Symeodonis, G. (2002), *The Effect of Competition Cartel Policy and the Evolution of Strategy and Structure in British Industry*, MIT Press.

Vasconcelos, H. (2006), "Endogenous mergers in endogenous sunk cost industries," *International Journal of Industrial Organization*, vol. 24, 227–250.

Walsh, P. and Whelan, P. (2002), "Portfolio Effects and Firms Size Distribution: Carbonate Soft Drinks," *The Economic and Social Review*, vol. 33, 43–54.

Wooldridge, J. (1997), "Quasi-maximum Likelihood Methods for Count Data," in "Handbook of Applied Econometrics Vol II," Oxford; Blackwells, Pesaran, H. and Schimdt, P. (eds), pages –.

Yen, S. and Adamowicz, W. (1993), "Statistical Properties of Welfare Measures from Count-Data Models of Recreation Demand," *Review of Agricultural Economics*, vol. 15, 203–215.

Yi, S. (1997), "Stable Coalition Structure with Externalities," *Game and Economic Behavior*, vol. 20, 201–237.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Nbanks	34.598	19.096	8	86	2673
C1	24.869	8.153	12.963	80.672	2673
CV	1.626	0.419	0.691	2.648	2673
DENS	0.033	0.044	0.004	0.264	2673
K	0.002	0.002	0	0.012	2673
P	0.06	0.019	0.022	0.104	2673
NB_OWN	0.001	0.002	0	0.043	2673
NB_COMP	0.045	0.048	0.002	0.231	2673
NB_OWN_OUT	0.005	0.01	0	0.093	2673
NB_COMP_OUT	0.177	0.156	0.002	0.631	2673

Table 2: Estimation results: Zero Truncated Poisson
Dependent variable: Numbers of Banks - equation (3.3)

Variable	**(1)**	**(2)**	**(3)**
NB_OWN	9.242***	9.080***	9.043***
	(2.062)	(1.967)	(1.744)
NB_COMP	8.190***	8.482***	8.404***
	(1.344)	(1.242)	(1.204)
NB_OWN_OUT	-2.699***	-2.122***	-0.874
	(0.623)	(0.566)	(0.594)
NB_COMP_OUT	-0.049	0.004	-0.011
	(0.182)	(0.216)	(0.205)
CV	0.167	0.228	0.231*
	(0.151)	(0.140)	(0.134)
K	-141.197***	-126.401***	-124.354***
	(48.777)	(41.179)	(39.865)
P	9.768***	12.335***	11.658***
	(2.876)	(2.900)	(2.737)
DENS	3.859***	3.163***	3.135***
	(1.322)	(1.105)	(1.065)
VA_pct	1.627	0.572	0.677
	(2.158)	(1.741)	(1.675)
REGIONE1		-0.204*	-0.198
		(0.123)	(0.121)
REGIONE2		-0.248**	-0.229*
		(0.123)	(0.118)
TYPE1			0.106***
			(0.031)
TYPE2			0.032***
			(0.012)
Constant	2.338***	2.237***	2.231***
	(0.219)	(0.200)	(0.191)
ll	-9490.302	-9347.755	-9266.454
N	2673	2673	2673
chi2	1051.747	1067.701	1097.830
p	0.000	0.000	0.000

* p<0.10, ** p<0.05, *** p<0.01

Table 3: Estimation results: Zero Truncated Negative Binomial
Dependent variable: Numbers of Banks

Variable	**(1)**	**(2)**	**(3)**
NB_OWN	9.227***	8.933***	9.042***
	(2.278)	(2.194)	(1.890)
NB_COMP	8.564***	8.615***	8.527***
	(1.428)	(1.279)	(1.236)
NB_OWN_OUT	-2.559***	-2.121***	-0.843
	(0.601)	(0.562)	(0.600)
NB_COMP_OUT	-0.032	0.026	0.010
	(0.170)	(0.206)	(0.195)
CV	0.163	0.219	0.222*
	(0.142)	(0.137)	(0.132)
K	-141.481***	-124.063***	-121.415***
	(48.087)	(42.405)	(40.864)
P	8.761***	11.128***	10.515***
	(2.610)	(2.885)	(2.751)
DENS	3.729***	3.081***	3.045***
	(1.294)	(1.161)	(1.115)
VA_pct	1.479	0.557	0.647
	(2.352)	(1.846)	(1.762)
REGIONE1		-0.177	-0.173
		(0.124)	(0.121)
REGIONE2		-0.209*	-0.193
		(0.123)	(0.118)
TYPE1			0.114***
			(0.032)
TYPE2			0.036***
			(0.013)
Constant	2.392***	2.289***	2.278***
	(0.193)	(0.189)	(0.181)
lnalpha	-3.676***	-3.831***	-3.930***
	(0.371)	(0.425)	(0.422)
ll	-9196.711	-9138.896	-9087.824
N	2673	2673	2673
chi2	600.999	641.091	657.921
p	0.000	0.000	0.000

* p<0.10, ** p<0.05, *** p<0.01

Table 4: Estimation results: Marginal effects

Variable	(ztp)	(ztnb)	(Variable Mean)
NB_OWN	283.1834	283.0865	.0011403
	(52.357)	(56.701)	
NB_COMP	263.1779	266.9689	.0447364
	(36.787)	(38.366)	
NB_OWN_OUT	-27.37672	-26.40309	.005079
	(18.463)	(18.601)	
NB_COMP_OUT	-.3452477	.2979915	.1772281
	(6.431)	(6.110)	
CV	7.227972	6.952858	1.625817
	(4.238)	(4.167)	
K	-3894.202	-3801.284	.0018006
	(1238.163)	(1277.182)	
P	365.0903	329.2003	.0600712
	(87.430)	(87.960)	
DENS	98.17784	95.34141	.0327039
	(32.976)	(34.739)	
VA_pct	21.21327	20.25852	.0143725
	(52.418)	(55.120)	
REGIONE1 (d)	-6.113763	-5.35278	
	(3.697)	(3.708)	
REGIONE2 (d)	-6.920037	-5.862967	
	(3.473)	(3.508)	
TYPE1 (d)	3.423496	3.673721	
	(1.037)	(1.093)	
TYPE2 (d)	1.002621	1.144445	
	(0.392)	(0.417)	

(d)marginals for discrete change of dummy variable from 0 to 1

Table 5: Estimation results: QUANTILE REGRESSIONS

Variable	Quantile5%	Quantile25%	Quantile50%
1/log(POP)	51.416***	37.429***	34.027***
	(5.967)	(12.168)	(10.508)
REGIONE1	0.186	0.070	0.005
	(0.162)	(0.160)	(0.107)
REGIONE2	0.253	0.035	-0.025
	(0.180)	(0.172)	(0.114)
TYPE1	-0.041	-0.122	-0.204*
	(0.111)	(0.179)	(0.114)
TYPE2	-0.046	0.094	-0.005
	(0.116)	(0.200)	(0.129)
Constant	-5.737***	-4.225***	-3.644***
	(0.455)	(0.949)	(0.817)
C_∞ REGION1	0.0039	0.0154	0.0256
	(0.0021)	(.0144)	(0.0205)
C_∞ REGION2	0.0041	0.0149	0.0249
	(0.0020)	(0.0142)	(0.0198)
C_∞ REGION3	0.00321	(0.0144)	0.0255
	(0.0015)	(0.0135)	(.0203)
N	103	103	103

*p<0.10,**p<0.05,***p<0.01

Table 6: Estimation results: QUANTILE REGRESSIONS

Variable	Quantile5%	Quantile25%	Quantile50%
1/log(banks)	7.556*	7.078	2.947
	(4.440)	(5.069)	(3.001)
REGIONE1	0.428***	0.232	0.029
	(0.099)	(0.199)	(0.100)
REGIONE2	0.437***	0.196	-0.022
	(0.132)	(0.202)	(0.102)
TYPE1	0.011	-0.108	-0.198**
	(0.186)	(0.203)	(0.097)
TYPE2	0.057	-0.019	-0.054
	(0.089)	(0.237)	(0.116)
Constant	-3.123***	-2.498***	-1.500***
	(0.719)	(0.818)	(0.489)
C_∞ REGION1	0.0633	0.0940	0.1824**
	(0.0455)	(0.0646)	(0.0730)
C_∞ REGION2	0.0638	0.0909	0.1791***
	(0.0441)	(0.0655)	(0.0679)
C_∞ REGION3	0.0422	0.0760	0.1868***
	(0.0290)	(0.0574)	(0.0683)
N	99	99	99

*p<0.10,**p<0.05,***p<0.01

Table 7: Estimation results: OLS - Reduced form
Dependent variable: NB_OWN

Variable	Coefficient
NB_OWN_03	0.9798***
	(0.008)
NB_COMP	0.0003*
	(0.000)
NB_OWN_OUT	0.0002
	(0.001)
NB_COMP_OUT	0.0001
	(0.000)
CV	-0.0000
	(0.000)
K	-0.0025
	(0.009)
P	0.0002
	(0.000)
DENS	0.0002
	(0.000)
VA_pct	0.0003
	(0.000)
TYPE1	0.0000
	(0.000)
TYPE2	0.0001***
	(0.000)
REGION1	-0.0000
	(0.000)
REGION2	0.0000
	(0.000)
Constant	0.0000
	(0.000)
R^2	.9883225
N	2414.000
p	0.000

* p<0.10, ** p<0.05, *** p<0.01

Table 8: Estimation results: Augmented Zero Truncated Negative Binomial
Dependent variable: Numbers of Banks

Variable	**(1)**
NB_OWN	9.170***
	(1.923)
NB_COMP	8.556***
	(1.255)
NB_OWN_OUT	-0.591
	(0.587)
NB_COMP_OUT	0.006
	(0.203)
CV	0.225*
	(0.135)
K	-121.924***
	(41.176)
P	10.713***
	(2.718)
DENS	3.087***
	(1.128)
VA_pct	0.613
	(1.770)
TYPE1	0.121***
	(0.032)
TYPE2	0.034**
	(0.014)
REGION1	-0.181
	(0.124)
REGION2	-0.201*
	(0.120)
vhat	-2.433
	(8.790)
Constant	2.260***
	(0.186)
	-3.852***
	(0.404)
ll	-8245.051
N	2414.000
chi2	666.529
p	0.000

* p<0.10, ** p<0.05, *** p<0.01

Table 9: Estimation results: GLM
As the dependent variable C1 is a proportion, it is to use generalized linear model with family binomial family, logistic link function. Standard errors scaled using square root of Pearson X2-based dispersion

Variable		
1/log(POP)	43.956***	
	(12.273)	
REGIONE1	-0.173	-0.145
	(0.117)	(0.136)
REGIONE2	-0.227*	-0.188
	(0.124)	(0.139)
TYPE1	-0.234*	-0.263*
	(0.132)	(0.138)
TYPE2	0.040	-0.088
	(0.145)	(0.160)
1/log(banks)		-0.848
		(4.023)
Constant	-4.221***	-0.686
	(0.957)	(0.656)
C_∞ REGION1	0.0122	0.3034**
	(.0115)	(0.1256)
C_∞ REGION2	0.0116	0.2943**
	(0.1098)	(0.1283)
C_∞ REGION3	0.01450	0.3349**
	(0.0144)	(0.1462)
N	103	99

*p<0.10,**p<0.05,***p<0.01

Chapter 2

Unit Roots and the Dynamics of Market Shares: an analysis using Italian Banking micro-panel

Abstract

The paper proposes the use of panel data unit root tests to assess market share instability in order to have (preliminary) indications of the industry dynamic. The idea is to consider the movements in market shares not only as element of the market structure but rather reflecting competitors' conduct. If shares are mean-reverting, the firm actions have a temporary effect on shares only. On the other hand, if they are evolving, as signaled by the presence of unit roots, the gain in shares respect with the competitors could be long-term. To illustrate the potential of unit roots tests, I consider an application to the Italian retail banking industry.

Keywords: turbulence, cross-section dependence.
JEL Classification: C23; D40.

2.1 Introduction

In order to get an understanding of the dynamic of an industry, a first step could be to examine whether the market shares are stationary or evolving. If shares are mean-reverting, the firm actions only have a temporary effect on shares. On the contrary, if they are evolving, as signaled by the presence of unit roots, the gain (or loss) in shares respect with the competitors could be long-term. In the first case, it is reasonable to infer that the industry is rather stable - or mature - where actors reached positions difficult to overcome. In the second case, instead, the possibility for a competitor to become permanently a leader (or to loose the leadership) could be a signal that the industry experienced the displacement of existing technology by alternative ones and/or the displacement of existing products by new and superior substitutes. In other words, by considering the movements in market shares not only as elements of the market structure but rather reflecting competitors' conduct that arise from the market (Asplund and Nocke (2006); Caves (1998); Matraves and Rondi (2007); Sutton (2004); Uchida and Cook (2005)), unit root tests could be a way to empirically test the influence of industry characteristics on the degree of turbulence (Davies and Geroski (1997); Sutton (1997)).

An important characteristic of market share data sets is the logical consistency requirement in market share models. In fact, market shares are bounded between 0 and 1 and they sum to unity. This relationship must be taken into account if one want to study all the actors in the market[1]. Another possibility is to consider only few actors in the market.

According to this latter procedure, this paper proposes the use of micro-panel data unit root models to assess market share instability in the Italian Banking Industry for a sample of firms made of the first 5 banks in each province. On the one hand, the assessment of the competitive conditions of the Italian banking industry is of interest since the industry has known a marked consolidation process along with a remarkable deregulation pro-

[1] An interesting approach is presented in the work of Franses et al. (2001). They exploit the consistency requirement to apply the Johansen test, relying on a system-based test rather than a single equation test. In addition, the fact that the data are bounded from below and above renders a deterministic model implausible.

cess since the beginning of the nineties. To the best of my knowledge, this paper is the first to apply this methodology to test banking competition. On the other hand, given the well-known low power of conventional unit root tests when applied to single short time series, panel unit root tests can be fruitfully employed in analysis of firms or industries that rely on micro-panels, where the time dimension may be of limited length but observed across several units. One of the main advantages of panel unit root tests is that their asymptotic distribution is standard normal. This is in contrast to individual time series unit roots which are non-standard normal asymptotic distributions. However, these tests are not exempt from critics. In particular, few tests consider the possibility of cross-section correlation and spillovers amongst countries, regions or provinces (Baltagi et al. (2007)). In this regard, Pesaran (2004) suggests a test for cross-sectional dependence and way of getting rid of it by augmenting the usual ADF regression with lagged cross-sectional mean and its first-difference to capture the cross-sectional dependence that arises through a single factor model. Other important aspects concerning panel unit roots are related to their asymptotic behavior under the two dimension of the panel, N and T and their requirement for a balanced panel (no missing data for any i not t).

Clearly, the use of panel unit root tests can only offer (preliminary) indications of the dynamic in the industry. As any other statistical test, there is a risk of incorrect inference but it could be minimized by properly selecting the test in relation to the main features of the dataset used. In any case, results must be supported by other qualitative - and if it is feasible - quantitative evidence. However, the existence of dynamic in the positions of the first 5 banks - as signalled by the presence of unit roots in the market shares - suggests that the Italian retail banking industry experienced overtime a movement towards higher level of competition. In particular, in the same spirit of Kim et al. (2003), a dynamic in market shares could be interpreted as an indirect signal for a reduction of switching costs that make easier to consumers to move to different banks and, consequently, for banks to acquire new customers.

The structure of the paper is as follows. The next section briefly introduces the data and the main features of the industry under investigation. Section 3 presents the model for micro-panels, whereas in section 4 panel unit roots tests are computed for the first 5 banks in every Italian province. Section 5 provides the Pesaran's test of cross-section dependence. The following section, taking into account these results, computes the unit root test proposed by Pesaran which deals with the cross-section dependence. The conclusions are presented in the final section.

2.2 Characteristics and construction of the dataset

A peculiarity of the Italian banking industry is the presence of different territorial dynamics (Colombo and Turati (2004); Guiso et al. (2004, 2006)). In particular, the retail Italian Banking Industry can be viewed as made of a large number of local markets corresponding to different geographical locations. In each one of these submarkets, there are several branches of different banks competing against each other. The Italian territory is divided into 20 regions and 103 provinces, which are geographical units close to US counties. In accordance with the Italian Antitrust Authority, the presumption is that the province is the relevant market.

Given the widespread differences in local economic conditions and their influence on the competition process, I will focus on local markets by measuring market shares at the provincial level using data on branches as proxies for the market share of individual (or group of) banks. The reasons behind the choice to compute market shares relying on this variable are various. First of all, the number of branches (or branch density) is commonly used in the empirical literature on local banking competition (see, for instance, Degryse and Ongena (2005)). Secondly, it captures the dimension of banking competition that has been more heavily affected by the deregulation process. Since March 1990, the establishment of new branches has been completely liberalized. The number of branches increased steadily, up to 32.337 in 2007, as well as the number of people served by each branch, 47 per 100.000 inhabitants in 2004 (compared to

59 EU mean)[2]. In addition, this measure is made freely available, without any break, for a long period of time by the Italian Central Bank (Banca d'Italia) [3].

Hence, the (unbalanced) dataset is composed of 103 Italian provinces. For each province I computed the market shares for the first 5 individual banks (or group of banks) from the year 1993 to 2006. The majority of Italian banks do not belong to any groups.

Table 1: Market Share Summary statistics

Year	Mean	Std. Dev.	Min.	Max.
1993	0.136	0.11	0.017	0.707
1994	0.136	0.11	0.016	0.683
1995	0.138	0.107	0.016	0.672
1996	0.137	0.105	0.016	0.646
1997	0.138	0.106	0.023	0.643
1998	0.140	0.112	0.023	0.849
1999	0.145	0.111	0.025	0.843
2000	0.143	0.108	0.024	0.836
2001	0.139	0.102	0.024	0.829
2002	0.139	0.1	0.024	0.807
2003	0.137	0.099	0.024	0.808
2004	0.135	0.098	0.024	0.808
2005	0.133	0.095	0.024	0.807
2006	0.131	0.093	0.025	0.802
N		515		

In this work, the unit of observation is the bank (or group of banks) in each province and year. This means that I treat the share of the same banks (or group of banks) in a different province as pertaining to a different bank. In order to consider the dynamics at (higher) regional level, tests will also be performed by grouping the different provinces according to different

[2]Beginning in the 1980s, the Italian Banking system underwent a series of reforms aimed at increasing the competition in the market through liberalizing branching and easing the geographical restrictions on lending. In fact, the opening of new branches had been regulated by the branch distribution plan, issued every four years. The last distribution plan was issued in 1986.

[3]`http://siotec.bancaditalia.it/sportelli/main.do?function=language\&language=ita`.

macro-region: North, Centre and South. As Guiso et al. (2004, 2006) showed, while there is a considerable variation in the degree of banking competition across local markets, the North-Centre/South divide is a clear feature of the Italian banking Industry.

Table (1) and (2) report the summary statistics of the market share of the first five banks in the sample. A closer look at these tables seem to reveal a stable pattern over time and the North-Centre/South divide. However, they may indicate little since there might be an intensive switching among banks' positions and a greater variability at local level. The challenge of the proposed methodology is to find out the underlying dynamic of banks at provincial level.

Table 2: Market Share Summary statistics

	Mean	**Std. Dev.**	**Min.**	**Max.**	**N**
NATIONAL	0.138	0.104	0.016	0.849	7210
NORTH	0.135	0.092	0.016	0.525	2590
CENTRE	0.14	0.102	0.026	0.548	2100
SOUTH	0.138	0.117	0.024	0.849	2520

North: Friuli-Venezia Giulia, Liguria, Lombardia, Piemonte, Trentino Alto-Adige,Veneto
Centre:Emilia Romagna, Lazio, Marche, Toscana, Umbria
South: Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sardegna, Sicilia

2.3 Tests for unit roots: the model

In this paper I consider micro-panel data models where the cross-section dimension is much larger than the time series dimension. Reviews of the literature on dynamic micro-panels are provided in Baltagi (2005) and Arellano (2003). For a general survey of the literature about unit root tests see Breitung and Pesaran (2006).

Let s_{it} be the market share of bank i in period t in each province. The model could be represented by a dynamic AR(1) panel data model allowing for heterogeneity in the intercept but not in the autoregressive parameter

$$
\begin{aligned}
s_{i0} &= \delta_0 + \delta_1 \eta_i + v_{i0} \\
s_{it} &= \alpha_i s_{i,t-1} + u_{it} \\
u_{it} &= (1 - \alpha_i)\eta_i + v_{it}
\end{aligned}
\tag{2.1}
$$

where $\alpha_1 = \cdots = \alpha_N = \alpha$ for each $i = 1 \ldots N$, $t = 2 \ldots T$, and where N is large and T is fixed. The series have a unit root (or are integrated of order one) if $\alpha_i = 1$ and are stationary if $\alpha_i < 1$. A test for the presence of a unit root in the panel is presented by the null hypothesis $H_0 : \alpha = 1$ in equation (2.1).

In case of independence across firms, the error term satisfies

$$
E(\eta_i) = 0, \quad E(v_{it}) = 0 \tag{2.2}
$$

for $i = 1, \ldots, N$ and $t = 2, \ldots, T$ and

$$
E(v_{it} v_{is}) = 0 \tag{2.3}
$$

for $i = 1, \ldots, N$ and $t \neq s$.

In the literature two types of panel unit root tests can be distinguished, dependent on the alternative under consideration. The first type of test considers a homogeneous alternative, i.e $H_0 = \alpha_1 = \cdots = \alpha_N = \alpha < 1$. An example is Levin et al. (2002). The idea of this approach is to perform a pooled Dickey-Fuller (DF) test with the residuals. The second type of test allow for heterogeneity of all parameters. Im et al. (2003) criticize the assumption of common root under the alternative and they require $|\alpha_i| < 1$ for a sufficiently large number of units. Consequently in this case, it is natural to perform N tests individually and to average over individual DF statistics.

In that model there are two sources of persistency. One is the autoregressive mechanism, which is the same for all cross-section units, and the other is the unobserved individual-specific term. The unit root hypothesis can be considered as an extreme case where all the presistency is caused by the autoregressive mechanism. In this context, the time dimension of the panel dataset is an important issue to look at, as well as the specification of the initial value.

In this paper, two distinct regression-based test procedures will be proposed: one based on a simple OLS regression of market shares on their lagged values; the second test, proposed by Breitung and Meyer (1994), specifies the regression in terms of deviations from initial conditions and it is therefore more powerful if the variance of individual effects is high. These simple t-tests based on least-squared estimators, which are consistent only under the unit root null, are shown to have good size properties and at least as high power as test based on GMM and ML estimators (Bond et al. (2005)). It is known that instrumental variable and GMM procedures provide consistent estimate of dynamic coefficients in cases where pooled least squares are inconsistent (Arellano (2003), Phillips and Sul (2007)). However, these procedures are also known to suffer bias and weak instrumentation problem when the dynamic coefficient α_i is close to unity.

However, these tests are not exempt from critics. In particular, they assume cross-section independence. Hence, the Pesaran's test for cross-section independence will be computed. The test can in fact be applied to a wide range of panel data model, including panel with short time dimension. As that test evidenced the presence of cross-section dependence, the panel unit root test allowing for cross-section dependence proposed by Pesaran will be also computed.

2.3.1 OLS

As Bond et al. (2005); Madsen (2003); Hall (2002) have shown, the t-test based on the OLS levels estimator performs much better than other estimators (GMM, FD, WG,..) in micro panels, that is when T is very small in comparison with N. Both simulation and asymptotic analysis have demonstrated that the OLS estimator has the highest power to reject alternative that are close to the null hypothesis that $\alpha = 1$.

Because the number of periods is small, properties of the initial condition are also relevant. Madsen (2003) shows that the asymptotic power of the OLS test under the alternative differs depending on the assumption made about the initial value. In particular, the advantage of using OLS is expected to be high when the initial value are such that the time-series process become covariance stationary, even for value of α close to unity[4]. In the other cases, when the initial values are such that the time-series become mean stationary and when the variation in the individual-specific terms is high, the highest power can be obtained using a t-test for the least squares estimator in the transformed model proposed by Breitung and Meyer (1994). As emerged from the previous works, these two tests must be considered jointly.

Under the null $H_0 : \alpha = 1$, the OLS estimator of α in model 2.1 is consistent. The t-test based on OLS estimator is

$$t_{OLS} = \frac{\hat{\alpha}_{OLS} - 1}{\sqrt{\hat{Var}(\hat{\alpha}_{OLS})}} \tag{2.4}$$

where

$$\hat{Var}(\hat{\alpha}_{OLS}) = (s'_{-1}s_{-1})^{-1}\left(\sum_{i}^{N} s'_{i,-1}e_i e'_i s_{i,-1}\right)(s'_{-1}s_{-1})^{-1} \tag{2.5}$$

with $e_i = s_i - s_{i,1}\hat{\alpha}_{OLS}$, $s_i = (s_{i,2}, \ldots, s_{i,T})'$, $s_{i,-1} = (s_{i,1}, \ldots, s_{i,T-1})'$, and $s_{-1} = (s'_{1,-1}, \ldots, s'_{N,-1})'$. Under the null, $\alpha = 1$, t_{OLS} has an asymptotic standard normal distribution as $N \to \infty$.

Under the alternative, the OLS estimator is biased upwards, more so when the variance of η_{it} is large relative to the variance of v_{it}. The power of this test will therefore depend on the magnitude of $Var(\eta_i)/Var(v_{it})$ (Bond et al. (2005)).

Breitung and Meyer (1994) suggest an alternative estimation approach which involves deducting the first observation s_{i0} for each firm from the right hand side of equation (2.1). The estimable model becomes

$$s_{it} - s_{i0} = \alpha(s_{i,t-1} - s_{i,0}) + \epsilon_{it} \tag{2.6}$$

$$\tilde{s}_t = \tilde{s}_{t-1} + \epsilon_{it} \quad t = 3, \ldots, T \tag{2.7}$$

[4]Mean stationarity (constant first moment) requires $\alpha_i < 1$ and $\delta_0 = 0$ and $\delta_1 = 1$. The covariance stationarity (constant first and second moments) in addition requires homeschedatisticity over time of the v_{it} shocks (i.e. $var(v_{it}) = \sigma^2_{vi}$ for $i = 1, \ldots, N$) and that $var(v_{i0}) = \sigma_{vi}/(1 - \alpha_i)^2$.

where $\epsilon = v_{it} - (1-\alpha)(s_{i0} - \eta_i)$. Again, the OLS estimator is consistent when $\alpha = 1$ and upward biased under the alternative. Breitung and Meyer however showed that the bias is $\alpha + \frac{1-\alpha}{2}$. That means that the power of the test, contrary to the previous case, is not affected by the individual-specific term, that is by the term $\frac{Var(\eta_i)}{Var(v_{it})}$. For long T panels, none of these could be applied, since the asymptotic distribution tends to a DF: so it is necessary to combine N DF/ADF tests as in Im et al. (2003).

2.3.2 A test for Cross Section Dependence

Pesaran (2004), Baltagi et al. (2007) show that there can be considerable size distortions in panel when the hypothesis of cross section independence is violated and the specification exhibits, for example, spatial error correlation.

When N is small and the time dimension T is sufficiently large, the cross section correlation can be modeled using seemingly unrelated regression (SURE), and traditional times series techniques - such the Lagrange Multiplier (LM) of Breusch and Pagan - can be applied[5]. However in cases where N is large, standard techniques are not applicable. Another approach, used in the literature of spatial statistics, measures the extent of cross dependence by means of a spatial matrix.

Pesaran (2004) proposes instead a simple diagnostic test that neither requires any *a priori* specification of a connection matrix nor suffers of panel data model limitations. It is therefore applicable in a variety of contexts, including stationarity dynamic and unit-root heterogeneous panels with short T and large N. The test, in all its various formulation, is based on simple averages of pair-wise correlation coefficients of OLS residuals from individual regressions.

[5]For example, Chu et al. (2007) used the panel SURADF tests to investigate Gibrat's law of proportionate effects for 48 electronic firms in Taiwan. Panel SURADF tests handle cross-sectional dependence across firms and, at the same time, investigate a separate unit-root null hypothesis for each and every individual panel member, identifying how many and which series in the panel are stationary process

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right) \tag{2.8}$$

where

$$\hat{\rho_{ij}} = \frac{\sum_{t=1}^{T} \hat{v_{it}} \hat{v_{jt}}}{\left(\sum_{t=1}^{T} \hat{v_{it}}^2\right)^{1/2} \left(\sum_{t=1}^{T} \hat{v_{jt}}^2\right)^{1/2}} \tag{2.9}$$

Unlike the LM statistic, the CD statistic has exactly mean at zero for fixed value of T and N, under a wider range of panel data model, and it is shown to have a standard normal distribution, assuming that the errors are symmetrically distributed, v_{it} are $i.i.d.(0, 1)$. In addition, it can be applied to unbalanced panels. In this last case, equation (2.10) can be modified by

$$CD = \sqrt{\frac{2}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sqrt{T_{ij}} \hat{\rho}_{ij} \right) \tag{2.10}$$

where $T_{ij} = \sum \mathbf{1}_{T_i \cap T_j}$ (the number of common time series observations between units i and j) and

$$\hat{\rho_{ij}} = \frac{\sum_{t \in T_{ij}} (\hat{v_{it}} - \bar{\hat{v_{it}}})(\hat{v_{jt}} - \bar{\hat{v_{jt}}})}{\left(\sum_{t=1}^{T} (\hat{v_{it}} - \bar{\hat{v_{it}}})^2\right)^{1/2} \left(\sum_{t=1}^{T} (\hat{v_{jt}} - \bar{\hat{v_{jt}}}))^2\right)^{1/2}} \tag{2.11}$$

with $\bar{\hat{v_{it}}} = \sum_{t \in T_{ij}} \hat{u_{ij}} / T_{ij}$. Finally, in cases where the cross section units can be ordered *a priori*, as with spatial observations, the CD test can be generalized capturing the spatial pattern too (Pesaran (2004)).

2.4 The Italian case

As previously stated, the Italian banking industry is of interest since it experienced a deregulation process during the nineties that lead, among other things, to liberalized entry and easier procedures to open new branches.

Therefore, it is reasonable to expect a movement towards higher level of competition, both at National and macro-regional level.

Let's start with the two simple t-tests based on OLS regressions.

Table 3: OLS:
Dependent variable: Firm shares

Variable	NATIONAL	NORTH	CENTRE	SOUTH
t*	2.91	3.17	0.65	1.06
p-value	0.004	0.002	0.517	0.290
with time dummies				
t*	4.8	6.93	6.2	4.8
p-value	0.000	0.000	0.000	0.000
N	5758	2086	1695	1977

*Note: the t-statistic for H_0: $\beta = 1$ against $H_1 : \beta < 1$

Results relying on constrained estimations - imposing restrictions on the constant being equal or greater than zero - are substantially identical.

Table 4: Breitung and Meyer 1994:
Dependent variable: Firm shares

Variable	NATIONAL	NORTH	CENTRE	SOUTH
t*	3.19	2.5	0.31	2.9
p-value	0.001	0.012	0.750	0.004
N	5758	2086	1695	1977

*Note: the t-statistic for H_0: $\beta = 1$ against H_1: $\beta < 1$.

As table (3) and (4) show, the null hypothesis $H_0 : \alpha = 1$ cannot always be rejected and the series seem to present unit roots, especially in the market-share related to the provinces in the Centre/South of Italy. These results disappear when it is controlled for common shocks (captured by year dummies). If the hypothesis of unit root cannot be rejected, it means that the positions of the main banks in the market could be displaced permanently by other actors, and there is evidence for a shift towards higher degree of competition.

2.4.1 Cross Section Dependence in Dynamics Panels

In the previous section, as it is typically assumed in panel data models, disturbances are treated as cross sectionally independent. To check if the panel at hand is characterized by cross-section dependence, the CD test is applied[6]. Two specification are used: one with residuals of homogeneous regression and one with residuals from N individual regressions from model (2.1). In both cases, the tests draw also on the residuals of the specification with/without the intercept and the observations are grouped again according to macro-regional classification (North, Center and South)[7]. The correlation are computed over the common set of observations for i and j. As it is already noted, the OLS estimates of the constant, $(1-\alpha_i)\eta_i$, and slope α_i for the individual series are biased when T is small. And that bias could be substantial for value of α near unity. However, the CD TEST advanced by Pesaran (2004) is valid for all values of α in model 2.1, including unity. The main reason lies in the fact that despite the sample bias of the parameters estimates, the OLS have exactly mean zero even for a fixed T, so long as the errors are symmetrically distributed. The main limitation of Pesaran's test relies in its pair wise construction since it could be possible that pair wise correlation compensate each others, summing to zero. While allowing for different value of α_i and cross-section correlation, we still assume that in model (2.1) v_{it} are serially uncorrelated with zero mean. Due to computational reasons, I restrict the analysis to the balanced panel made of 20% of the observations. Table (5) reports results for the cross section dependence test developed by Pesaran. In both cases, allowing or not for heterogeneity, there is evidence for cross section dependence in all the macro-region considered.

[6] For this test I built a STATA command csdar.ado relying on xtcsd.ado as developed by De Hoyos and Sarafidis (2006).

[7] Results in table (5) refer to estimation without intercept. Those with intercept are analogous and are not reported.

Table 5: Pesaran's Test of Cross Section Independence

	NATIONAL	NORTH	CENTRE	SOUTH
Residuals from a regression*				
CD Statistic	6.875	50.370	44.920	52.694
P-value	0.000	0.000	0.000	0.000
Residuals from N regression*				
CD Statistic	5.106	50.228	44.334	52.864
P-value	0.000	0.000	0.000	0.000

* without intercept

2.5 Panel Unit Root Tests for Cross Sectionally Dependent Panels

Overall, the outcome of the preceding tests clearly indicates the presence of cross section dependence amongst units.

Pesaran builds on the assumption that the error terms v_{it} of equation (2.1) follow a single common factor structure

$$v_{it} = \lambda_i f_t + \epsilon_{it} \tag{2.12}$$

The common factor is assumed to be stationary and to impact the cross-section with a fraction determined by the individual specific factor loading λ_i. Because of the common factor, cross section dependence arises and can be approximated by the cross-section mean $\bar{s}_t = \frac{1}{N}\sum_{i=1}^{N} s_{it}$. As usual, the ϵ_{it} are assumed to be $i.i.d$ across i and t with zero mean and variance σ_i^2, and $E(\epsilon_{it})^4 < \infty$. Furthermore, ϵ_{it}, f_t and λ_i are mutually independently distributed for all i.

Pesaran proposes the following augmented Dickey-Fuller regression:

$$\Delta s_{it} = c_i + \rho_i s_{i,t-1} + \beta_i \bar{s}_{t-1} + \sum_{j=1}^{p} \gamma_{ij} \Delta s_{i,t-j} + \sum_{j=0}^{p} \gamma_{ij} \Delta \bar{s}_{i,t-j} + \epsilon_{it} \tag{2.13}$$

where, as usual in the univariate case, lagged first-differences on both s_i and $\bar{s}_i$ are added in order to take into account also for possible correlation in the error term. Either individually or in a combined fashion, the t-value of ρ_i can be used to test the presence of unit roots. In the first case, the statistic is called cross-sectionally augmented Dickey-Fuller ($CADF_i$) while in the second case the statistic is constructed as

$$CIPS = \frac{1}{N}\sum_{i}^{N} CADF_i \tag{2.14}$$

It is called CIPS since it resembles the IPS statistic (Im et al. (2003)). In the case were T is fixed, to ensure the $CADF$ statistics do not depend on the nuisance parameters, Pesaran (2003) suggests to apply the test to the deviations of the variable from initial cross-section mean.

Table 6: Pesaran's Test for Unit Root

	NATIONAL	**NORTH**	**CENTRE**	**SOUTH**
CADF (0 lag)				
t	-1.563	-2.020	-2.067	-2.086
z[t-bar]	1.178	-1.309	-2.067	-2.086
P-value	0.881	0.095	0.023	0.084
CADF (1 lag)				
t	-1.333	-2.708	-1.590	-1.262
z[t-bar]	3.024	-1.560	0.744	1.644
P-value	0.999	0.059	0.772	0.950

Table (6) shows the Pesaran's test for unit root computed using Italian Banking dataset. Due to the presence of the lagged level of the cross sectional average, the limiting distribution of the CADF statistics and the CIPS statistic does not follow a standard Dickey-Fuller distribution. However, Peseran provides critical values based on simulations for the CADF and CIPS-distributions for three cases (no intercept and no trend, intercept only, intercept and trend). As results from this test suggests, cross-section dependence does matter. When controlling for it, the series exhibit the presence of unit roots. That means that there is a dynamic in the positions of the main competitors in the market, and there exist the possibility for the main actors to be displaced by competitors. The question to be addressed now is what are the factors which drive the results. It seems likely that more than one factor play a role. Merger and acquisition activity, the presence of scale economies, and the role of regulation each appear to have had a role. As suggested by Kim et al. (2003), movements in market shares can also be used to infer (and measure) switching costs. By making it costly for consumers to change bank - and consequently more difficult for a new bank to acquire new clients, switching costs tend to limit entry as well as shuffle in market shares. To this end, other investigations are required. However, from this simple analysis is possible to infer that the Italian Banking industry experienced movement towards higher levels of competition. In particular, these results match those on the local level competition of Guiso et al. (2004). By looking at the long-term effect of the regulatory restriction, they found that, after the deregulation process, there was a catching up of the areas (especially the provinces in the South) where the banking market was less competitive during the regulation period. So, the presence of unit root in the market share data of this macro-region is consistent with their analysis.

2.6 Conclusions

The paper proposed the use of unit root tests in the setting of micro panel data sets to assess market share instability in order to get an understanding of the competitive condition in an industry. Using Italian Banking micro-panel, this study empirically tests the presence of unit roots in the series of market shares of the first five banks in each province. The presence of unit roots in the market share data could be interpreted, in fact, as signal of an industry that experienced the displacement of the leading bank by its competitors. On the other hand, if share turn out to be mean reverting, it is reasonable to conclude that the industry is rather stable and competitors reached positions difficult to overcome.

Two simple t-tests based on least squares estimators, which are consistent only under the unit root null, have been proposed. Those tests are shown to have good size properties and at least as high power as tests based on GMM and ML estimators. According to those tests, the hypothesis of unit root tests cannot always be rejected for all the subgroups considered in the analysis.

However, these tests do not consider the possibility of cross-section correlation amongst units. To check if the panel at hand is characterized by cross-section dependence, the Pesaran's cross-section dependence test was applied. The Pesaran's statistics clearly indicated the presence of cross-section dependence. As a consequence, the ADF regression proposed by Pesaran was applied. In that case, results strongly confirmed the presence of unit root tests.

The kind of exercises performed in this article could only offer (preliminary) indications of the dynamic of market shares in an industry. To individuate which factors drive the results, of course, other analyses are required. Nevertheless, as this simple application to the Italian banking case shows, panel unit root tests are useful and versatile tools that, combined with an institutional knowledge of the industry under investigation, could offer interesting insights on the industry competitive process.

References

Arellano, M. (2003), *Panel Data Econometrics*, Oxford: Oxford University Press.

Asplund, M. and Nocke, V. (2006), "Firm Turnover in Imperfectly Competitive Markets," *Review of Economic Studies*, vol. 73, 295–237.

Baltagi, B., Bresson, G. and Pirrotte, A. (2007), "Panel Unit Root Tests and Spatial Dependence," .

Baltagi, B. (2005), *Econometric Analysis of Panel Data - Third Edition*, Wiley.

Bond, S., Nauges, C. and Windmeijer, F. (2005), "Unit Roots: Identification and Testing in Micro Panels," *Cemmap Working Paper No. 07/05.*

Breitung, J. and Meyer, W. (1994), "Testing for unit roots in panel data: are wages on different bargaining levels cointegrated?" *Applied Economics*, vol. 26, 353–361.

Breitung, J. and Pesaran, H. (2006), "Unit Roots and Cointegration in Panels," in "The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice," Kluwer Academic Publishers, ch.8, forthcoming, in L. Mátyás and P. Sevestre (eds.), pages –.

Caves, R. (1998), "Industrial Organization and New Findings on the Turnover and Mobility of Firms," *Journal of Economic and Literature*, vol. 36, 1947–1982.

Chu, H., Sher, P. and Yeh, M. (2007), “Revisiting Gibrat’s law using panel SURADF tests,” *Applied Economics Letters*, vol. 1, 1–7.

Colombo, L. and Turati, G. (2004), “The role of local real and financial variables in banking industry consolidation: the case of Italy,” Working Paper 80, Department of Economics and Finance.

Davies, S. and Geroski, P. (1997), “Change in Concentration, Turbulence, and the Dynamics of Market Shares,” *The Review of Economics and Statistics*, vol. 79, 383–391.

De Hoyos, R. and Sarafidis, V. (2006), “Testing for Cross-sectional Dependence in Panel Data Models,” *The Stata Journal*, vol. 6, 482–498.

Degryse, H. and Ongena, S. (2005), “Distance, Lending Relationship and Competition,” *Journal of Finance*, 231–266.

Franses, P.H., Srinivasan, S. and Boswijk, P. (2001), “Testing Unit Roots in Market Shares,” *Marketing Letters*, vol. 12, 351–364.

Guiso, L., Sapienza, P. and Zingales, L. (2004), “Does Local Financial Development Matter,” *The Quarterly Journal of Economics*, 929–969.

Guiso, L., Sapienza, P. and Zingales, L. (2006), “The Cost of Banking Regulation,” *CEPR Discussion Paper, No 5864.*

Hall, B. and Mairesse, J. (2002), “Testing for unit roots in panel data: An exploration using real and simulated data,” .

Im, K., S.C., A. and Shin, Y. (2003), “Testing for Unit Roots in Heterogeneous Panels,” *Journal of Econometrics*, vol. 115, 53–74.

Kim, M., Kliger, D. and Vale, B. (2003), “Estimating switching costs: the case of banking,” *Journal of Financial Intermediation*, vol. 12, 25–56.

Levin, A., Lin, F. and Chu, C. (2002), “Unit root tests in panel data: Asymptotic and finite-sample properties,” *Journal of Econometrics*, vol. 108, 1–24.

Madsen, E. (2003), “Unit Root Inference in Panel Data Models where Time-Series Dimension is Fixed: a Comparison of Different Tests,” *CAM Working Paper No. 13*.

Matraves, C. and Rondi, L. (2007), “Product Differentiation, Industry Concentration and Market Share Turbolence,” *International Journal of Economic and Business*, vol. 14, 37–57.

Pesaran, M.H. (2004), “General Diagnostic Tests for Cross Section Dependence in Panels,” *CESifo Working Paper No. 1229*.

Phillips, P. and Sul, D. (2007), “Bias in Dynamic Panel Estimation with Fixed Effects, incidental trends and cross section dependence,” *Journal of Econometrics*, vol. 137, 162–188.

Sutton, J. (1997), “Gibrat’s Legacy,” *Journal of Economic Literature*, vol. 35, 40–59.

Sutton, J. (2004), “Market Share Dynamics and the Persistence of Leadership’ Debate,” *London School of Economics Working Paper*.

Uchida, Y. and Cook, P. (2005), “Innovation and Market Structure in the Manifacturing Sector: an Application of Linear Feedback Models,” *Oxford Bullettin of Economics and Statistics*, 557–580.

Chapter 3

Relationship Lending and Firm Innovativeness: New Empirical Evidence

Abstract

The aim of this study is to investigate the effects of relationship lending on firm innovativeness, disentangling the impact of bank ties on the discovery phase from that in the introduction phase of new technologies. Results suggest that for small and low-tech firms, banks do not carry out a sophisticated intervention at the stage of development of new technologies and, rather, they play their traditional role in financing investments of otherwise financially constrained firms. On the contrary, relationship banks do play an important role even in the discovery phase for high-tech firms.

Keywords: Credit relationship, external financing, bank competition.
JEL Classification: C34, G21, O31.

3.1 Introduction

As the fourth Community Innovation Survey highlighted, for the majority of Italian firms the main obstacle inhibiting innovation is still represented by financial factors. Firm's financial need, however, is not constant and varies in relation to firm characteristics as well as project phases. For example, during the so-called *seed* phase, the financial need to carry out a feasibility study is rather low, whereas it is high in the *start-up* phase, when the project has to be implemented. During the *early growth* stage, instead, a firm requires considerable fundings in order to market its innovative products. Lacking the visibility of more established firms, young and small firms are likely to suffer even more for the financing of their investments because of asymmetric information problems (Berger and Udell (1998)). Moreover, the different phases of a project are characterized by different degree of risk.

Relying on a panel of Italian manufacturing firms, the main objective of the following empirical investigation is to examine the effects of relationship lending on firm innovativeness identifying two phases of the innovation process: the 'discovery phase', which roughly corresponds to the seed and start-up phases, and the 'introduction phase', which in the spirit of this paper is similar to the early growth phase.

As Italy has strongly relied on relationship banking to finance investments, focusing on Italian firms allows to isolate the role of bank-firm relationship in fostering innovation. In particular, Italian banks seem better suited to financing innovation embodied in physical capital rather than R&D investment. As it is documented in Ughetto (2007), there is a striking difference in the share of bank loans as a source of funds for fixed capital compared with that one for R&D projects. In the same line, Herrera and Minetti (2007) suggest that relationship banks do not carry out a sophisticated intervention at the stage of assessment and development of new technologies and they, rather, play their traditional role in financing investments of otherwise financially constrained firms.

In order to disentangle the effects of bank ties on the discovery phase from those in the phase of introduction and adoption of new technologies, the following analysis proceed in two steps, first by measuring firm *propensity to innovate*, and, then, by estimating firm *intensity to innovate* in the introduction phase, as measured by the percentage of new products in total sales. From the econometric perspective, it means to adopt a generalised tobit model which tries to account for the fact that firms are either innovative or not, and, for those that are innovative, the extent to which they are so (Mohnen et al. (2006)). This strategy has several advantages. Firstly, it allows to distinguish between different levers that bank can use to influence innovation, among which the provision of funds is of course the most important. Distinguishing between invention and introduction of new technologies is also important in the light of firm innovation patterns, as Italian firms tend to absorb innovations from outside than in carrying out research. Finally, it allows to control for selectivity problems.

As the empirical determinants of relationship lending have been already investigated in the literature (i.e., Elsas (2005)), I will only give a recall of its key elements so as to focus on the reasons why bank ties should affect firm innovative capacity.

To get a complete picture, since the dataset also offers indications on the other type of financing, I will take into account other sources of finance available to the firms, as well as the role of public incentives. Since banks are by far the most important source of external finance in Italy, it is reasonable to expect the internal sources to play a crucial role in financing innovation.

Another important peculiarity of the Italian banking system is its delimitation in local areas, corresponding to 103 provinces which are geographical units close to US counties (Colombo and Turati (2004); Guiso et al. (2004, 2006)). The geographical segmentation is relevant in order to identify the level at which competition indicators, which are important control variables in the present analysis, have to be computed.

Closely related to this analysis are the works of Benfratello et al. (2007) and Herrera and Minetti (2007). However, it differentiates from them for various reasons. First of all, the more recent dataset used. This work will in fact rely also on the last Capitalia survey (the ninth). It also departs from Benfratello et al. (2007) for having a deeper look at the effects on innovation of specific bank-firm ties (such as the duration), distinguishing the discovery from the introduction phase, instead of focusing on the level of financial development. With respect to the work of Herrera and Minetti (2007), the methodology adopted is substantially different. These authors in fact investigated the possibility of endogeneity of the relationship variables while estimating the probability of introducing innovation. In this work, the method of estimation should account for this problem, considering in addition other important variables that in their work have been neglected. In particular, the effects of intensity of fixed capital and R&D Investment, as well as the role of internal source of financing in the probability of being innovative.

The paper is structured in the following way. Next section gives an overview of the literature while section 3, after a brief recall of the empirical determinants of relationship lending, explores its possible links with firm innovativeness. Section 4 presents the dataset and the main descriptive statistics on the degree of firm innovativeness. Section 5 presents the model and the results by distinguishing the introduction and discovery phases of innovation. On the contrary, section 6 presents a deeper analysis for the discovery phase only. The final section summarizes the paper.

3.2 Literature Review

A lively macro-economic debate on the role of financial architecture in fostering innovation and technology is the one on the bank-based versus market-based system (i.e.,Carlin and Mayer (2003), Levine (2002)). Are bank-based systems at the advantage in processing information particularly relevant for firms' incentive to innovate? The available evidence is rather mixed but findings suggest that market-based system do not dominate bank-based system and vice-versa in all times. However, knowledge-intensive industries, with soft, hard-to-monitor complex activities seem to get on better in bank based financial systems (Tadesse (2007)).

Even though researchers have argued theoretically, and tested empirically, that there is a link between finance and innovation, there is still little in the extant micro-economic literature about the functions of the various sources of funding in the different phases of innovation (O' Sullivan (2004)). The main contribution of this paper is to make another step in this direction, enhancing our understanding of the role played by banks ties in the phases of invention and introduction of new technologies.

In fact, this paper relates to two, somehow separated, strands of empirical literature. The first comprehends articles on the economics of innovation. During the past decade, a number of countries in Europe have implemented enterprise-based surveys of innovative activity (i.e., Community Innovation Survey (CIS). At the same time, important progress has been made in modelling appropriate econometric methods for innovation survey data. Hall and Mairasse (2006) provide an interesting review of the empirical studies on innovation.

The second strand mainly relates to works investigating bank-firm relationship. Since there is a vast literature, I only refer to works related to the Italian banking system. For a review of the literature see Elyasiani and Goldberg (2004) and Degryse and Ongena (2008). Some of these works investigate the credit access for firms belonging to industrial districts. Relying on the ninth Capitalia survey Ughetto (2006) and Rotondi(2005) show that firms in industrial districts are less likely to be credit rationed. In particular, Alessandrini et al. (2008) evidence that the incidence of relationship lending for firms in industrial districts is not significantly different from the average. Ferri and Messori (2000) show that arm's length patterns prevail in the Northwest, the area of oldest industrialization with larger banks and firms, whereas relationship banking patterns prevail in the rest of the country, populated to a larger extent by smaller banks and firms.

Someways in-between, there are the works of Benfratello et al. (2007) and Herrera and Minetti (2007) which, instead, stress - at micro level - the role of Italian banks in fostering innovation. Benfratello et al. (2007) find strong evidence that banking development has a significant and important impact effect on process innovation and a weaker effect on product innovation. In addition, they find that banking development has lessened the severity of financing constraints faced by small firms. Herrera and Minetti (2007) test the impact on innovation of the information of the main bank

- proxied by the duration of credit relationship. They observe that firms with longer credit relationship have higher probability to innovate. Furthermore, the length of the relationship seem to foster the acquisition of new technologies rather than internal research. Using a large panel of US companies, Atanassov et al. (2007) explore the relationship between arm's length financing and innovation taking patents as a measure of innovative output. They found that firms that relied more on arm's length financing are associated with a larger number of patents. They also conclude that this correlation is mainly driven by innovative firms choosing their capital structure. Relying on firm-level data from a survey conducted in Finland, but looking instead at the role of public policy, the work of Hyytinen and Toivanen (2005) provides evidence that capital-market imperfections delay innovation, and government funding disproportionately helps firms in industries that are more dependent on external finance. They used as a measure of firm innovativeness the level of R&D expenditure.

3.3 Relationship lending and Firm innovativeness

In this section, in order to identify the main variables to be used in the empirical analysis, first, I will recall the key elements of relationship lending. Then, I will investigate the reasons why relationship lending should affect the phases of innovation. For a detailed description of the empirical determinants of relationship lending see Elsas (2005), whereas for an analytical survey on the effects of relationship lending on the pricing of loans, as well as its effect on the degree of competition, see Freixas (2005).

3.3.1 Empirical determinants of relationship lending

Relationship lending represents the informational privilege that a bank accumulates over time by establishing close ties with its borrower (Ongena and Smith (2001)). Reflecting the idea that long tenure depicts the relationship intensity, the most commonly proxy for relationship lending is the *duration* of a bank-borrower relationship. The exclusivity of bank relationship is also regarded as an indicator of close ties between the bank and the borrower. In this regards, the *number of bank relationships* should capture the possibility for bank to realize the economic benefit associated with the relationship. A negative correlation between the number of banks and the development of relationship lending is reasonable. Finally, a higher debt *financing share* should increase the likelihood of relationship lending.

3.3.2 Relationship lending and innovation

In Italy relationship lending has always been a way to channel funds to productive investments. In fact, despite its development, the stock market does not play a crucial role, while specialized financiers play a marginal role. In 2004, in the comparison between the Italian and the European venture capital industry - in term of venture capital and private equity instruments over GDP - Italy ranked 12th, with all other large European economies ranking well above. In addition, the Italian venture capital industry is focused on later-stage investments: on average, in 2004, early stage financing represented only 2% of total investments in Italy compared with 6.4% in Europe (Gregoriou et al. (2006)). Banks, in particular, turned out to be better suited to finance innovation embodied in physical capital rather than technological progress (Ughetto (2007)). As Italian firms typically do not receive external equity, internal equity finance (auto-financing/cashflow) represented an important source of innovation financing as well. Capitalia survey shows that, in 2001-2003, for 83% of firms auto-financing still represents the main source to finance innovation, followed by 10% of firms relying more on public incentives, and 5% on banks loans.

Different theoretical arguments point out that investment in R&D activities is different from investment in capital goods. First, R&D project may not be easily understood by outsiders and create large intangible assets which cannot be used as a collateral (Hall (2002)). In addition, expected returns of R&D are uncertain and difficult to estimate. Finally, as suggested by Bhattacharya and Chiesa (1995), firms may be reluctant to finance externally their R&D project for strategic reasons.

In which way then, relationship lending affects firm innovative capacity? Can it mitigate firm resort to internal finance? What makes a relationship lender special?

In addition to have a direct effect on the quantity of R&D and investment spending, banks may also affect the nature of the selected project, the quality of internal inputs as well as their effectiveness in generating innovation. In particular, their ability to offer multi-period contracts, which are much more effective than one-shot contracts (i.e. transactions) in extracting information, may be helpful in the allocation process and the mechanisms that allow firms to make commitments of resources to innovative activities, notwithstanding the challenges of doing so. As Boot (2000) argued, relationship banking goes beyond lending and includes other services as well. In this regard, relationship lending leaves room for flexibility and discretion allowing the utilization of non-contractable information and addressing contractual features that are possibly unique. Furthermore, the firm can disclose information to the bank without worrying about it spilling over competitors.

On the other hand, close and durable relationship may involve inefficiencies related to the hold-up and soft-budget-constraint problems. The hold-up problem refers to the possibility that relationship bank may extract rents thus causing inefficient choice investment (see von Thadden (1995), and for a review Allen and Carletti (2008)), whereas the soft-budget-constraint problem concerns the bank's incentive to refinance some of the ex-post inefficient projects (Dewatripont and Roland (2000)).

3.4 Data description

The data used in this work are obtained by the two most recent waves - the 8th and the 9th - of the comprehensive survey on Italian manufacturing firms carried out by Capitalia (and previously by Mediocredito Centrale) every three years[1]. These surveys are conducted through questionnaires, administered to a representative sample of manufacturing firms within the national borders. Questionnaires collected information over the previous

[1] See "Indagine sulle imprese manifatturiere" http://www.capitalia.it/pages/studi02b.htm.

three years (1998-2000 and 2001-2003) and, for the majority of the firms, are supplemented with standard balance sheet data. The 8th and the 9th survey include respectively 4289 and 4497 firms. To broaden the sample period of the analysis, I merged these two waves and obtained a reduced sample of 2097 firms. This sample includes only those firms exiting in both surveys and therefore with potentially complete observations over the 1998-2003 period. I further excluded firms with incomplete information or with extreme values. I will progressively use the panel structure of the data in order to check and address the endogeneity problems.

Based on this sample, tables (1) and (2) report the population percentages (and standard errors) of firms with either product or process innovation. The most important information is the increasing percentage of innovative firms, across size and sectors, over the period considered (the only exception is the % of firms with more than 500 employes doing process innovation). These higher percentages reflect the higher number of firms doing R&D. As table (3) shows, particularly in high-tech industries, the majority of firms are envolved in R&D activities. This is even more visible for larger firms where this percentage reached 92% in high-tech sectors[2]. Table (4) reports the (population) mean of the variables measuring relationship lending for the period 2001-2003. There are not significant differences in the duration of the relationship with the main bank, in the bank main share, and in the number of lending banks between small low-tech and high-tech firms, as well as for large low-tech and high-tech firms. There are significant differences when comparing these values according to the size variable. Interesting to note, however, is that there are no significant differences for small and large high-tech firms in the mean value of the variables related to the main bank (duration and share).

[2]Firms where classified as in:

- low-tech sectors: textile, wood, food, plastic, paper, coke, non metallic and nec (not elsewhere classified).

- high-tech sectors: vehicles, machinery and chemicals

3.5 The empirical model and results

I adopt a generalized (Type 2) Tobit model consisting of two equations, where the first one is a probit equation determining whether a firm innovates or not ("propensity to innovate"), and the second one is a linear regression (the Tobit equation or "intensity to innovate") explaining how much the firm innovates (Mohnen et al. (2006)). I will measure firm innovative propensity by means of new processes and new products introduced into the market, whereas the firm innovation intensity can be measured by the share of innovative sales in total sales. Contrary to other type of surveys (i.e, Community Innovation Survey - CIS), it is not possible to distinguish between innovative sales corresponding to products new to the firm but possibly known to the market, which can be considered imitations of products already produced by other competitors, and those corresponding to products only new to the market, which can be regarded as true innovations.

Denoting by y_{1i} the binary variable indicating if firm i is an innovating firm - that is, a dummy variable indicating whether the firm either has introduced at least one product or process innovation - I can write

$$y_{1i} = \begin{cases} = 1 & if \quad y^*_{1i} > 0 \\ = 0 & if \quad y^*_{1i} \leq 0 \end{cases} \tag{3.1}$$

where $y^*_{1i} = x_{1i}b_1 + u_{1i}$ is a latent variable that represents the incentives to innovate. x_{1i} is a vector of explanatory variables, b_1 is a vector of parameters to be estimated, and u_{1i} is a random error term, which includes the effect of left-out omitted variables. As explanatory variables x_{1i}, in addition to the amount of resources spent on R&D per employee ($[IE]$) and fixed capital per employee($[INVEST]$), I use an industry dummy ($[HIGH\_TECH]$), a size variable ($[LOGSIZE]$), and a dummy for listed company ($[LISTED]$). The industry dummy ($HIGH\_TECH]$) captures technological opportunity conditions, industry-targeted innovation policies, and high-tech specific differential demand growth effect. Size - measured by the log of the number of employees - reflects access to finance, scale economies and difference in the organization of work (Mohnen et al. (2006)). In order to account for the fact that young firms grow faster

(Klomp 1996), I add a dummy for firms that are less than three years old ([$YOUNG$]). It could be valuable to include a dummy also for firms that underwent structural change ([$M\&As$]) during the period of the analysis and for firms operating in international markets ([$INTERN.\ COMP$]).

As the main objective of my investigation is to control how relationship lending affect firm innovativeness, I estimate the probability to be innovative controlling for relationship lending including in the explanatory set, x_{1i}, variables representing

- the share of the main bank: [$BANK\_SHARE$]
- the duration of the relationship: [$LENGTH$]
- the number of bank lenders: [NUM_BANKS]

Finally, to account for the possibility to have access to other sources of funding, I include in the regressors a dummy variable, [FIN_INSTR], for firms that relied on innovative financial instruments, such as financial bills, project finance, or private bond. The second equation of the Tobit (type 2) model is specified in terms of a second latent variable y^*_{2i} which is equal to the actual share of innovative sales y_{2i}, if the firm is innovative (i.e, $y^*_{1i} > 0$). Since the share of innovative sales is bounded by 0 and 1, it is preferable to perform a logit transformation of the data and express this second equation in terms of the latent logit-share variable $z^*_{2i} = ln(y^*_{2i}/(1 - y^*_{2i}))$ which vary from $-\infty$ to $+\infty$. Thus I can write our second equation as

$$z_{2i} = \begin{cases} = z^*_{2i} & if \quad y^*_{1i} > 0 \\ = undefined & if \quad y^*_{1i} \leq 0 \end{cases} \tag{3.2}$$

or equivalently

$$y_{2i} = \begin{cases} = e^{z^*_{2i}}/(1 + e^{z^*_{2i}}) & if \quad y^*_{1i} > 0 \\ = 0 & if \quad y^*_{1i} \leq 0 \end{cases} \tag{3.3}$$

where $z^*_{2i} = x_{2i}b_2 + u_{2i}$.

x_2 is a vector of explanatory variables, b_2 is a vector of parameters to be estimated and $u_{2i} > 0$ is an error term reflecting omitted variables. Since I have data on sale growth for the majority of the firms in the panel, I decide to exploit the panel structure of the data in order to exclude the variable past sales growth ($[g\_sales_{[t-1]}]$) from the explanatory variables I have in x_{2i}, and to include it in x_{1i}. This variable in fact can be a determining factor of innovation, as reflecting stronger demand and easier internal and external access to finance. There are a lot of missing values in the variables of interest. For example, many firms do not indicate the amount of resources spent on R&D. The final sample is made of 1258 observations for the period 2001-2003. I also present some results using as exclusionary variable $[rationed_{[t-1]}]$, a dummy variable which is equal to 1 if the firm answered to be credit rationed in the previous survey. Results are substantially equal, even though this variable resulted more significant in some specifications. However, $g\_sales_{[t-1]}$ seems more reliable since it based on balance sheet data - instead of being determined by firm assessment - and it offers indications on the role of internal sources [3].

Assuming that u_1 and u_2 are bivariate normal with zero mean, and $\sigma_{u1} = 1$, I can estimate the model as a generalized Tobit (type 2) model using STATA Heckman procedure for survey analysis. Preliminary results for the model without considering any financial variables are reported in table (5). Table (6) reports results for the basic model relying on $rationed_{[t-1]}$ instead of $g\_sales_{[t-1]}$ as exclusionary variable. Those preliminary results suggest the plausibility of the model, and the significance of the ρ coefficient indicates selection problems in the intensity innovation equation. Results for traditional regressors are in line with the literature. Firms with higher spending on R&D and fixed investments are those most likely to introduce an innovation. Larger and listed firms, especially in high-tech industries, are also more likely to be innovative and to have a higher percentage of sales stemming from innovative products. International agreements on production as well as public incentives also positively affect firm capacity to innovate.

In table (7) results for the model controlling for financial variables are

[3]More precisely, firms are defined to be credit rationed if answer yes to all the following question: *1.* whether at the current market interest rate they wish a larger amount of credit; *2.* whether they would be willing to accept a small increase in the interest rate charged in order to obtain more credit 3. whether they have applied for credit but have been turned down by the financial intermediary.

reported. These results evidence that relationship variables do matter in explaining firm innovative capacity: the variables accounting for the share of the main bank and the number of lending banks are jointly highly significant, both in the intensity and propensity to innovate equation. However, theoretical and empirical works suggest that the market for SME finance is imperfect (see, for example, Alessandrini et al. (2007); Carpenter and Petersen (2002); Hyytinen and Toivanen (2005)): the opportunity cost of investments (the marginal cost of capital schedule) is higher for small firms (upward-sloping curve). That means, SMEs that are in need of (external) capital are more likely to pursue some innovations and positively affected by long-term relationship with some banks. To account for this possibility, and to control how it will affect the role of bank relationship in fostering innovation, table (7) also reports results for relationship variables interacted with a dummy variable for SMALL firms. A higher share of the major lending bank will have a positive effect on the capacity of small firms to translate innovation into a greater percentage of firm sales stemming from innovative products. This result is in line with those of Herrera and Minetti (2007) which found that relationship banks do not carry out a sophisticated intervention at the stage of assessment and development of new technologies. On the other hand, contrary to what have been found by Herrera and Minetti (2007), longer relationship may have counter positive effect on firm capacity to innovate. However, this result is not highly significant. The overall effect on both the capacity and intensity to innovate is significant at 10% level. Moreover, these results are robust and reinforced if $[SMALL]$ is replaced by $[LOGSIZE]$ as in the baseline regression, and if the exclusionary variable $[g\_sales[t-1]]$ is replaced by a dummy variable for firms being credit rationed in the previous survey (see table (8)).

3.5.1 Relationship Lending and Measure of Dependence on External Finance

Relying on the same set of variables used in the previous section, and by further exploiting the panel structure of the data, in this section I will estimate the previous model by identifying industries' *technological demand* *for external finance.* The reason for bringing into the picture this variable is related to the necessity to control for some specific industry features which may affect both the firm capacity to innovate and the role of bank ties. In order to do that, following Hyytinen and Toivanen (2005) and differently from Benfratello et al. (2007), I will compute my own measure for external dependence for Italian manufacturing firms, amending the Rajan and Zingales (1998)'s methodology (RZ). The main assumption of RZ is that there are technological reasons why some industries rely more on external finance than others (i.e, gestation periods of products, the initial project scale, the cash harvest period). It seems important therefore to look how these 'intrinsic' industry features may affect bank ties, and ultimately firm capacity to innovate. It is reasonable to think that the effects of relationship lending variables should vary with the needs of external capital: the more firms are dependent on external finance, the stronger the ties with banks, and the higher the effects on firm innovativeness. However, it would be risky to assume that industry demands for external financing in Italy will be the same of large listed US firms. Shifting the focus to between industry differences, therefore, I will build a measure of external finance dependence using firm-level variables as collected during the eight Capitalia survey. As in Hyytinen and Toivanen (2005), I estimate the measure of external finance, using a financial planning model (called also the *percentage of sales* approach, see Demirgüç-Kunt and Maksimovic (2002)). This index, denoted by EFN, measures the proportion of firms whose annual growth rate of sales exceeds the maximum growth rate that can be financed if a firm relies only on its internal resources and maintain its dividend (see box 1).

Box 1: **External Financial Need** The percentage of sales model relates a firm growth rate to its need for external funds. The external financing need, EFN_t, at time t growing at g_t percent a year is given by

$$EFN_t = g_t Assets_t - (1 + g_t)Earnings_t * b_t$$

On the right hand side, the first term is the required investment for growing at g_t percent while the second term is the internally available capital for investment, taking b_t - the proportion of the firm earnings that are retained for reinvestment at time t - as given. Earnings are calculated after interest and taxes. I compute two estimates of each firm's attainable growth rate. The maximum growth rate that can be financed if a firm relies only on its internal resources and maintains its dividend, IG, is obtained by assuming that the firm retains all its earnings (that is $b_t = 1$), equating EFN_t to zero and solving for g_t

$$IG_t = ROA_t/(1 - ROA_t)$$

where ROA is the firm's return on assets. Thus, more profitable firms can find more resources internally. Then, I compare for each firm in the sample its actual growth rate with the rate, IG, defined above. Finally, I compute for each industry (according to NACE classification) the proportion of firms in financial needs, that is those firms whose mean actual growth rate is above the mean maximum attainable. To check the robustness of the measure, I also compute the same percentage, assuming that firms does not pay dividends and obtain just enough debt financing to maintain a constant ratio of total debt to assets (implicitly also a summing that the firm does not issue equity or increase leverage). Again, setting EFN_t to zero, and using the value of equity in place of total assets, the growth rate is now equal to $SG = ROE/(1 - ROE)$. These measure are conservative in three ways. First, each maximum growth assumes that a firm utilizes the unconstrained sources of finance no more intensively that it is currently doing. Second, firms with spare capacity do not need to invest and may grow at a faster rate than predicted by the model. Third, the financial planning model abstract from technical advancement that reduce the requirements for investment capital. Thus it may overstate the cost of growth and underestimate the maximum growth rate attainable using unconstrained sources of finance (Demirgüç-Kunt and Maksimovic (1998, 2002)).

Those firms whose actual mean growth rates are above the maximum one are assumed to be in need of external finance. The main advantage of computing this index is to (partially) control also for reverse causality. As pointed out by Herrera and Minetti (2007), measures such as the length of the relationship might be endogenous to the innovation process. The econometric specification chosen, already accounts for selectivity problem, as the significance level of ρ coefficient in the various specification indeed points out. Since the measure of external financial dependence is computed at industry level (NACE classifications), and using data on the eight survey, it is not affected by the current firm behaviour. The introduction of such a measure in interaction with relationship lending variables should therefore account for the possibility that firm with greater financial need will tend to have longer/strength ties with lending banks. Results are reported in table (9) using the ROE version of the EFN index (see Box 1). In the first two column, the variable EFN is simply added to the basic model. In the last two column of table (9), EFN is introduced also in interaction with relationship lending variables. From that analysis three points deserve a remark. First of all, and contrary to the previous analysis for small banks, the variables accounting for relationship lending turned out to be jointly significant at 1% level in the 'discovery phase', suggesting that banks play an important role in the process of discovering new products and processes in those sector that grow faster, that is, which are more in need of external finance. Secondly, these regressions play down the importance of the share of the main banks (although still positive for firm with higher EFN), being significant only at 10% level in both equations. Finally, the length of the relationship and the number of lending banks are now significant - respectively at 10% and 5% level - in the propensity to innovate equation. In particular, they have a negative effect on the probability of introducing an innovation in those sectors that are more in financial need. Results using the ROA version of the EFN (not reported) instead only confirms the (weak) explanatory power of the bank share variable. Even in this case, results are not affected if the exclusionary variable $[g\_sales[t-1]]$ is replaced by a dummy variable for firms being credit rationed in the previous survey.

To sum up, the analysis performed controlling also for financial need confirms the importance of the role of the main bank in fostering innovation and suggest some possible counter positive effects when this relationship becomes longer.

3.5.2 Bank competition and Innovation

The level of competition among banks represents a factor which may either strengthen or weaken firm ties with the bank. On the one hand, there are theories that argue that competition and relationships are incompatible since banks may not enjoy the possibilities to extract profit later on in the relationship (Petersen and Rajan (1995)). On the other hand, other theories argue that more competition may instead increase relationship lending, allowing banks to mitigate the effects of fiercer competition extracting higher rents (Boot (2000); Boot and Thakor (2000)). Empirical works indeed suggest that competition and relationship lending are not necessary inimical (Degryse and Ongena (2007), Elsas (2005)). In order to bring this element into the analysis, I will compute an index of banking competition at provincial level. The Italian territory is divided into 20 regions and 103 provinces, which are geographical units close to US counties. In accordance with the Italian Antitrust Authority, the presumption is that the province is the relevant market. More specifically, I will include in the regressors the number of bank branches per squared kilimoter [$BANK\_COMP$] in each province. The branch density represents the monopolistic power of each branch and could be considered as a proxy of the (inverse of) transportation costs. More branches in the same provinces means, for each consumer, a lower distance to cover to reach a branch, a weaker power exerted by bank branch and an overall higher degree of competition (Degryse and Ongena (2005)). $BANK\_COMP$ is a measure similar to the one propose by Benfratello et al. (2007), the number of branches per habitants[4]. In addition to that index, I also consider a traditional measure of market concentration, represented by sum of the market share of the first three banks [$CH3$]. Even in this case, I will focus on local markets by measuring market shares at the provincial level using data on branches as proxies for the market share of individual (or group of) banks. However, one must be cautious in interpreting this measure as a proxy of banking competition. As Claessens and Laeven (2004)' analy-

[4]This measure can be considered as a proxy for the (inverse of) queuing costs. The less the population served by each branch (or the higher the number of branches for each individual), the lower the cost met by the customers.

sis shows, variables describing the banking system structure may not be good summary statistics for bank competitive environment. Conversely, these authors found that more concentrated banking system face a greater degree of competition. In this case, $BANK\_COMP$ and $CH3$ are negatively correlated (-0.41599). Table (10) presents results adding the banking competition controls. Interestingly, for both those measures, there is a negative correlation with the firm innovativeness, in the introduction as well as in the discovery phase. However, $BANK\_COMP$ is not significant. The concentration index $CH3$ is instead significant at 10% level. This result suggests that less concentrated credit markets might foster innovation (Spagnolo (2004)). It is interesting to see how that variable interact with the external financial need. Results are reported in the last two column of table (10). The relationship is negative and significant in the introduction phase, whereas it is not possible to reject the hypothesis that the overall effect in the discovery phase is zero. Instead, no significant interactions have been found between relationship and competition variables (not reported).

3.6 Relationship lending in the discovery phase

Relying on the same set of (time-varying) variables used in the previous section, and by completely exploiting the panel structure of the data, in this final section I will focus on the effects of bank ties in the discovery phase only, in order to better control for endogeneity issues. Given that I only have two observations about the introduction of innovation (in the eight and ninth survey), it is not possible to fully address the endogeneity problems and to identify casual links. However, since one fundamental problem is to control for unobserved firm characteristics that are constant over time, the conditional logit model will work properly. Conditional logit models eliminate the firm specific effects, but only switchers (that is, firms that introduced an innovation in just one of the two sub-periods) contribute to the likelihood function. Therefore, I can rely on a restricted number of observations, as only around 40% of the sample is made up of swicthers and not all the explanatory variables are observed for all firms in both periods. I cannot indeed control for another potential source of endogeneity caused by technological shock that leads, for example, to an increase both in the probability of observing an innovation and in the research intensity (Parisi et al. (2006)). As in the previous analysis, relationship lending variables turned out to be significant in explaining the probability of introducing process or product innovations. In the model presented in table (11), they are jointly significant at 10% level. The most important result is again the role played by the share of the main bank. However, and coherently with the previous analysis, it is not possible to reject the hypthosis that the overall effect of small firms is equal to zero. Again, banks seems not play a crucial role in the discovery phase for small firms. In the second column of table (11), I reestimate the model using a dummy variable for R&D, in $[IE]$'s stead, the variable measuring the amount of resource spent in R&D per employee, since there are firms that have reported to do R&D but were not able to indicate how much they spent for this purpose. In the same way, $[INVE]$ - replacing $[INVEST]$ - is a dummy variable equals to 1 if a firm declared that has invested in fixed capital but was not able to indicate the amount. Results are substantially identical.

In addition, in table (12), I repeat the same analysis but distinguishing the effects for high-tech and low-tech firms. At 10% level, it is possible to reject the hypothesis that the variables representing bank ties are jointly equal to zero for high-tech firms, whereas the same it is not true for low-tech firms. Looking at high-tech firms, then, it seems that bank might play a crucial role also in the introduction phase. In particular, for high-tech firms, the length of the relationship turned out to be individually significant at 10% level. Those results are confirmed and reinforced in column (2).

These regressions confirm the importance of bank ties in affecting firm innovative capacity, even though this is particularly true for firms in high tech sectors. The variables representing bank relationship behave in both regressions alike. Then, in line with the results of Herrera and Minetti (2007), it is possible to conclude that for small and low-tech firms, banks do not carry out a sophisticated intervention at the stage of assessment and development of new technologies and they, rather, play their traditional role in financing investments of otherwise financially constrained firms. For high-tech firms, instead, banks do play an important role even in the discovery phase.

3.7 Conclusions

Using data on sample of Italian manufacturing firms, this study investigated the effects of relationship lending on firm innovativeness, disentangling the impact of bank ties on the discovery phase from that in the introduction phase and adoption of new technologies. As Schumpeter argued in his earliest writing on the microeconomics of innovation, banks are pivotal players in the innovation process and play a central role in real-sector innovation, not merely as a conduits for the movement of capital funds from saver to entrepreneur. In Italy, in particular, relationship lending has always been a way to channel funds to productive investments, since both the stock market and specialized financiers have played a marginal role. However, despite the current richness of enterprise-based survey on innovative activity, there is still little in the extant micro-economic literature about the different role of the various sources of funding in the introduction and invention of new technologies. On the contrary, at a macro-level, there is a lively debate on the role of financial architecture (bank-based versus market-based) in fostering innovation and technology.

Results from the present micro-econometric analysis suggest that for small firms and low-tech firms, banks do not carry out a sophisticated intervention at the development stage of the innovation. Similarly to other analysis, Italian banks appear to play their traditional role in financing investments of otherwise financially constrained firms. In particular, a higher share of the main lending bank has a positive impact on the capacity of small firms to translate the innovation into a greater percentage of firms sales stemming from innovative products. On the contrary, relationship banks turned out to play an important role even in discovery phase for firms in high-tech sectors. These results are robust after controlling for a measure of external financial dependence across sectors and the level of banking competition. In particular, even if some regressions downplayed its role, the share of the main bank proved to be a key variable. The length of the relationship and the number of lending banks exhibit instead negative and significant effects on the probability of introducing an innovation, especially in those sectors that are more in financial need.

References

Alessandrini, P., Presbitero, A.F. and Zazzaro, A. (2008), "Banks and Firms in Industrial Districts (Banche e imprese nei distretti industriali)," *SSRN eLibrary*.

Alessandrini, P., Zazzaro, A. and Presbitero, A. (2007), "Bank Size or Distance: What Hampers Innovation Adoption by SMEs ?" Working Papers 304, Universita' Politecnica delle Marche (I), Dipartimento di Economia.

Allen, F. and Carletti, E. (2008), "The Roles of Banks in Financial Systems," in "Oxford Handbook of Banking," forthcoming, Berger, A. and Molyneux, P. and Wilson, J.

Atanassov, J., Nanda, V.N. and Seru, A. (2007), "Finance and Innovation: The Case of Publicly Traded Firms," *SSRN eLibrary*.

Benfratello, L., Schiantarelli, F. and Sembenelli, A. (2007), "Banks and Innovation: Microeconometric Evidence on Italian Firms," Boston College Working Papers in Economics 631, Boston College Department of Economics.

Berger, A. and Udell, G. (1998), "The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle," *Journal of Banking & Finance*, vol. 22, 613–673.

Bhattacharya, S. and Chiesa, G. (1995), "Proprietary Information, Financial Intermediation, and Research Incentives," *Journal of Financial Intermediation*, vol. 4, 328–357.

Boot, A. (2000), "Relationship Banking: What do we know?" *Journal of Financial Intermediation*, vol. 9.

Boot, A. and Thakor, A. (2000), "Can Relationship Banking Survive Competition?" *Journal of Finance*, vol. 55, 679–713.

Carlin, W. and Mayer, C. (2003), "Finance, investment, and growth," *Journal of Financial Economics*, vol. 69, 191–226.

Carpenter, R. and Petersen, B. (2002), "Is The Growth Of Small Firms Constrained By Internal Finance?" *The Review of Economics and Statistics*, vol. 84, 298–309.

Claessens, S. and Laeven, L. (2004), "What Drives Bank Competition? Some International Evidence," *Journal of Money, Credit and Banking*, vol. 36, 563–583.

Colombo, L. and Turati, G. (2004), "The role of local real and financial variables in banking industry consolidation: the case of Italy," Working Paper 80, Department of Economics and Finance. 10

Degryse, H. and Ongena, S. (2005), "Distance, Lending Relationship and Competition," *Journal of Finance*, 231–266.

Degryse, H. and Ongena, S. (2007), "The Impact of Competition on Bank Orientation," *Journal of Financial Intermediation*, vol. 16, 399–424.

Degryse, H. and Ongena, S. (2008), "Competition and regulation in the banking sector: A review of the empirical evidence on the sources of bank rents„" in "Handbook of Financial Intermediation and Banking," Elsevier, Thakor A. and A. Boot (eds.), pages –.

Demirgüç-Kunt, A. and Maksimovic, V. (1998), "Law, Finance, and Firm Growth," *Journal of Finance*, vol. 53, 2107–2137.

Demirgüç-Kunt, A. and Maksimovic, V. (2002), "Funding growth in bank-based and market-based financial systems: evidence from firm-level data," *Journal of Financial Economics*, vol. 65, no. 3, 337–363.

Dewatripont, M. and Roland, G. (2000), “Soft Budget Constraints, Transition, and Industrial Change,” *Journal of Institutional and Theoretical Economics (JITE)*, vol. 127, 245–.

Elsas, R. (2005), “Empirical determinants of relationship lending,” *Journal of Financial Intermediation*, vol. 14, 32–57.

Elyasiani, E. and Goldberg, L. (2004), “Relationship lending: a survey of the literature,” *Journal of Economics and Business*, vol. 56, 315–330.

Ferri, G. and Messori, M. (2000), “Bank-firm relationships and allocative efficiency in Northeastern and Central Italy and in the South,” *Journal of Banking and Finance*, vol. 24, 1067–1095.

Freixas, X. (2005), “Deconstructing relationship banking,” *Investigaciones Economicas*, vol. 29, 3–31.

Gregoriou, G.N., Kooli, M. and Kraeussl, R. (2006), *Venture Capital in Europe: Quantitative Finance.*

Guiso, L., Sapienza, P. and Zingales, L. (2004), “Does Local Financial Development Matter?” *The Quarterly Journal of Economics*, vol. 119, 929–969. 10

Guiso, L., Sapienza, P. and Zingales, L. (2006), “The Cost of Banking Regulation,” CEPR Discussion Papers 5864, C.E.P.R. Discussion Papers. 10

Hall, B. (2002), “The Financing of Research and Development,” Department of Economics, Working Paper Series 1004, Department of Economics, Institute for Business and Economic Research, UC Berkeley.

Hall, B. and Mairasse, J. (2006), “Empirical Studies of Innovation in the Knowledge-driven Economy,” *Economics of Innovation and New Technology,*, vol. 15, 289–299.

Herrera, A.M. and Minetti, R. (2007), “Informed finance and technological change: Evidence from credit relationships,” *Journal of Financial Economics*, vol. 83, 223–269.

Hyytinen, A. and Toivanen, O. (2005), “Do financial constraints hold back innovation and growth?: Evidence on the role of public policy,” *Research Policy*, vol. 34, 1385–1403.

Levine, R. (2002), “Bank-Based or Market-Based Financial Systems: Which Is Better?” *Journal of Financial Intermediation*, vol. 11, 398–428.

Mohnen, P., Mairesse, J. and Dagenais, M. (2006), “Innovativity: A Comparison Across Seven European Countries,” *CIRANO Série Scientifique.*

O’ Sullivan, M. (2004), “Finance and Innovation,” in “Handbook of Innovation,” Oxford University Press, J. Fagerberg, and D. Mowery and R. Nelson (eds), pages 240–265.

Ongena, S. and Smith, D.C. (2001), “The duration of bank relationships,” *Journal of Financial Economics*, vol. 61, 449–475.

Parisi, M., Schiantarelli, F. and Sembenelli, A. (2006), “Productivity, Innovation and R&D: Micro Evidence for Italy,” *European Economic Review*, vol. 50, 2037–2061.

Petersen, M. and Rajan, R. (1995), “The Effect of Credit Market Competition on Lending Relationship,” *Quarterly Journal of Economics*, vol. 110, 406–443.

Rajan, R. and Zingales, L. (1998), “Financial Dependence and Growth,” *American Economic Review*, vol. 88, 559–86.

Spagnolo, G. (2004), “Debt as a (Credible) Collusive Device: ‘Everybody Happy But the Consumer’,” Working Paper in Economics and Finance 243, Stockholm School of Economics.

Tadesse, S. (2007), “Innovation, Information and Financial Architecture,” William Davidson Institute Working Papers Series 877, William Davidson Institute at the University of Michigan Stephen M. Ross Business School.

Ughetto, E. (2007), “Does Finance Matter for R&D Investment? New Evidence from a Panel of Italian Firms,” Innovation Studies Working Paper 10, Politecnico di Torino.

von Thadden, E. (1995), “Long-Term Contracts, Short-Term Investment and Monitoring,” *Review of Economic Studies*, vol. 62, 557–75.

.1 APPENDIX I

In this section the variables used in the regressions are described. They are obtained from the 8th and 9th survey on Italian manufacturing firms carried out by Capitalia every three years.

INNOVATION: dummy variable which takes the value 1 if the enterprise reports having introduced new production processes or products during 2001

Pct_SALES: Share of turnover in 2003 due to new products or process introduced during 2001-2003.

IE:Average total expenditure for internal and external R&D divided per employees over the period 2001-2003.

R&D: dummy variable which takes the value 1 if the firm reports to have done R&D during the period 2001-2003.

INVEST: Average gross investments in innovative tangible goods per employees over the period 2001-2003.

INVE: dummy variable which takes the value 1 if the firm reports to have invested in innovative tangible during the period 2001-2003.

YOUNG: dummy equal to 1 if the firms is less then three years old

SMALL: dummy equal to 1 if the firms has less than 50 employees.

M&As: dummy variable which takes the value 1 if the firm's was involved in merger and acquisition dealings.

INTERN._COMP: dummy variable which takes the value 1 if the enterprise's most significant market is international (outside EU).

INTERN._AGR.: dummy variable which takes the value 1 if the enterprise has developed technical agreement with firms operating in international markets (outside EU).

PATENTS_BOUGHT: dummy which takes the value 1 if the firms bought patents during the period 2001-2003.

PATENTS_SOLD: dummy which takes the value 1 if the firms sold patents during the period 2001-2003.

PUBLIC_INCEN.: dummy which takes the value 1 if the firms relied on public incentives during the period 2001-2003.

LISTED: dummy which takes the value 1 if the firm is listed in the stock market
LOGSIZE: the log of the average number of employees during the period. 2001-2003
BANK_SHARE: the share of the main bank
LENGTH:the duration of the relationship with the main bank
NUM_BANKS: the number of bank lenders
FIN._INSTR.: Dummy which takes the value 1 if the firm relied on innovative financial instruments during the period 2001-2003
BANK_COMP: number of branches per squared kilometer
CH3: market share of the first three banks
EFN: index of external financial dependence computed using firm-level variables as collected during the period 1998-2000
g_sales[t-1]: the turnover growth rate computed using variables as collected during the period 1998-2000
rationed[t-1]: dummy variable which takes the value 1 if the firm turned out be credit rationed during the period 1998-2000

Table 1: % of firms with a product innovation

FIRMS SIZE	LOW-TECH		HIGH-TECH	
(n° employees)	**1998-2000**	**2001-2003**	**1998-2000**	**2001-2003**
11-20	0.160 (0.014)	0.253 (0.027)	0.268 (0.017)	0.454 (0.034)
21-50	0.220 (0.015)	0.323 (0.032)	0.355 (0.016)	0.501 (0.027)
51-250	0.281 (0.030)	0.346 (0.039)	0.415 (0.015)	0.569 (0.022)
251-500	0.355 (0.064)	0.613 (0.076)	0.421 (0.046)	0.702 (0.050)
>500	0.598 (0.102)	0.671 (0.090)	0.391 (0.042)	0.426 (0.045)

() standard errors

Table 2: % of firms with a process innovation

FIRM SIZE	LOW-TECH		HIGH-TECH	
(n° employees)	**1998-2000**	**2001-2003**	**1998-2000**	**2001-2003**
11-20	0.287 (0.017)	0.321 (0.029)	0.287 (0.017)	0.297 (0.031)
21-50	0.364 (0.018)	0.397 (0.033)	0.371 (0.016)	0.417 (0.026)
51-250	0.490 (0.034)	0.497 (0.041)	0.494 (0.015)	0.508 (0.023)
251-500	0.460 (0.066)	0.660 (0.074)	0.498 (0.047)	0.614 (0.053)
>500	0.534 (0.102)	0.549 (0.099)	0.494 (0.043)	0.379 (0.044)

() standard errors

Table 3: % of firms doing R&D

FIRMS SIZE	**LOW-TECH**		**HIGH-TECH**	
(n° employees)	**1998-2000**	**2001-2003**	**1998-2000**	**2001-2003**
11-20	0.216 (0.016)	0.372 (0.030)	0.220 (0.016)	0.462 (0.034)
21-50	0.317 (0.017)	0.529 (0.034)	0.347 (0.016)	0.590 (0.026)
51-250	0.452 (0.034)	0.720 (0.037)	0.480(0.015)	0.704 (0.020)
251-500	0.632 (0.066)	0.879 (0.051)	0.561 (0.047)	0.783 (0.049)
>500	0.835 (0.079)	0.873 (0.069)	0.791 (0.044)	0.923 (0.031)

() standard errors

Table 4: Summary Statistics: Population Mean - Period 2001-2003

	BANK_SHARE	**LENGHT**	**NUM_BANKS**
LARGE & LT	30.3299	20.0506	7.0150
	(1.451)	(0.819)	(0.235)
LARGE & HT	32.1561	18.0048	6.8789
	(2.009)	(1.013)	(0.260)
SMALL & LT	35.3025	16.8680	4.1788
	(0.913)	(0.330)	(0.068)
SMALL & HT	34.0666	17.0709	4.1107
	(1.558)	(0.582)	(0.102)

Small firms: less than 50 employees

Table 5: Estimation results: HECKMAN BASE RESULTS
In the propensity equation the dependent variable is a dummy variable which takes value 1 if the firm has introduced at least one product or process innovation whereas in the intensity equation the dependent variable is a logit transformation of the actual share of innovative sales. The exclusionary variable is $g\_sales_{[t-1]}$

	Intensity Eq(2)		**Propensity** Eq(1)	
IE	0.2278***	(0.028)	0.1438***	(0.032)
INVEST	-0.0046	(0.010)	0.0125**	(0.006)
YOUNG	0.1046	(0.624)	0.0569	(0.235)
M&As	0.6525*	(0.336)	0.2164	(0.149)
INTERN._COMP	0.3084	(0.228)	0.1603	(0.104)
PATENTS_BOUGHT	0.0900	(0.610)	0.1722	(0.324)
PATENTS_SOLD	-1.8541**	(0.887)	-0.5622	(0.364)
INTERN._AGR.	0.3725	(0.328)	0.3026*	(0.171)
PUBLIC_INCEN.	0.5420**	(0.234)	0.2723***	(0.091)
LISTED	2.5741***	(0.870)	1.2456***	(0.372)
LOGSIZE	0.1748	(0.119)	0.2274***	(0.059)
HIGH_TECH	0.2377	(0.245)	0.1996**	(0.096)
Constant	-10.0413***	(1.982)	-3.9594***	(0.806)
$g\_sales_{[t-1]}$			0.0285	(0.084)
ρ	0.9237***	0.0234		
σ	2.4707***	0.19138		
ll	-35393.72			
N	584		1258	

*p<0.10,**p<0.05,***p<0.01
Note: Regressions include dummies for area

Table 6: Estimation results: HECKMAN BASE RESULTS
In the propensity equation the dependent variable is a dummy variable which takes value 1 if the firm has introduced at least one product or process innovation whereas in the intensity equation the dependent variable is a logit transformation of the actual share of innovative sales. The exclusionary variable is $rationed_{[t-1]}$

	Intesity Eq(2)		**Propensity** Eq(1)	
IE	0.2260***	(0.028)	0.1453***	(0.032)
INVEST	-0.0051	(0.010)	0.0130**	(0.006)
YOUNG	0.1013	(0.620)	0.0646	(0.236)
M&As	0.6440*	(0.334)	0.2224	(0.150)
INTERN._COMP.	0.3010	(0.226)	0.1545	(0.105)
PATENTS_BOUGHT	0.0851	(0.603)	0.1778	(0.326)
PATENTS_SOLD	-1.8268**	(0.873)	-0.5608	(0.364)
INTERN._AGR.	0.3658	(0.325)	0.3098*	(0.172)
PUBLIC_INCEN.	0.5229**	(0.234)	0.2797***	(0.092)
LISTED	2.5191***	(0.865)	1.2605***	(0.371)
LOGSIZE	0.1659	(0.119)	0.2351***	(0.059)
HIGH_TECH	0.2336	(0.245)	0.1974**	(0.097)
Constant	-9.8536***	(1.976)	-4.0201***	(0.803)
rationed$_{t-1}$			0.3398*	(0.194)
ρ	0.9196***	(0.0254)		
σ	2.4521***	0.1926		
ll	-35360.57			
N	584		1258	

*p<0.10,**p<0.05,***p<0.01
Note: Regressions include dummies for area

Table 7: Estimation results: HECKMAN ADDING FINANCIAL VARIABLES

	Intensity Eq(2)	**Propensity** Eq(1)	**Intensity** Eq(2)	**Propensity** Eq(1)
IE	0.2349***	0.1396***	0.2356***	0.1408***
	(0.030)	(0.030)	(0.031)	(0.031)
INVEST	-0.0064	0.0122*	-0.0064	0.0124**
	(0.011)	(0.006)	(0.010)	(0.006)
YOUNG	0.0339	0.0268	0.0831	0.0295
	(0.623)	(0.234)	(0.615)	(0.233)
M&As	0.6397*	0.2060	0.6583*	0.2181
	(0.350)	(0.151)	(0.348)	(0.152)
INTERN._COMP	0.3074	0.1652	0.3092	0.1706*
	(0.234)	(0.103)	(0.233)	(0.103)
PATENTS_BOUGHT	0.0478	0.1818	0.1366	0.2098
	(0.576)	(0.299)	(0.547)	(0.292)
PATENTS_SOLD	-1.7819**	-0.5457	-1.9887**	-0.5056
	(0.823)	(0.359)	(0.877)	(0.366)
INTERN._AGR.	0.3818	0.2922*	0.3833	0.2778*
	(0.341)	(0.164)	(0.334)	(0.164)
PUBLIC_INCEN.	0.5859**	0.2676***	0.5951**	0.2666***
	(0.235)	(0.090)	(0.235)	(0.091)
LISTED	2.4746***	1.2991***	2.7991***	1.2784***
	(0.875)	(0.375)	(0.815)	(0.356)
BANK_SHARE	0.0114***	0.0032**	-0.0021	-0.0013
	(0.004)	(0.002)	(0.006)	(0.003)
NUM_BANKS	0.0566	0.0306*	0.0822**	0.0347**
	(0.037)	(0.016)	(0.040)	(0.017)
LENGTH	-0.0133	0.0003	0.0216*	0.0027
	(0.011)	(0.004)	(0.013)	(0.006)
FIN._INSTR	-0.4861	0.2980	1.3071*	0.6892**
	(0.543)	(0.241)	(0.730)	(0.325)
SMALL	-0.3079	-0.2663***	0.3389	-0.2871
	(0.223)	(0.096)	(0.585)	(0.244)
HIGH_TECH	0.2260	0.2191**	0.2417	0.2239**
	(0.250)	(0.096)	(0.251)	(0.096)
Constant	-9.6042***	-3.3918***	-10.7604***	-3.3197***
	(1.882)	(0.788)	(1.787)	(0.746)
NUM_BANKSxSMALL			-0.0389	-0.0055
			(0.062)	(0.026)
BANK_SHARExSMALL			0.0143*	0.0048
			(0.007)	(0.003)
LENGTHxSMALL			-0.0414**	-0.0033
			(0.018)	(0.007)
FIN._INSTRxSMALL			-2.1136**	-0.4595
			(0.979)	(0.418)
past_sales$_{[t-1]}$		0.0604		0.0631
		(0.083)		(0.084)
ρ	0.9462***	(0.0162)	0.9463***	(0.0160)
σ	2.5408***	(0.1895)	2.5408***	(0.1895)
ll	-35236.53		-35159.27	
N	584	1258	584	1258

Table 8: Estimation results: HECKMAN ADDING FINANCIAL VARIABLES

	Intensity Eq(2)	**Propensity** Eq(1)	**Intensity** Eq(2)	**Propensity** Eq(1)
IE	0.2330***	0.1401***	0.2338***	0.1409***
	(0.030)	(0.031)	(0.031)	(0.031)
INVEST	-0.0067	0.0131**	-0.0070	0.0134**
	(0.010)	(0.006)	(0.010)	(0.006)
YOUNG	0.0237	0.0386	0.0374	0.0451
	(0.615)	(0.234)	(0.612)	(0.232)
M&As	0.6182*	0.1974	0.6241*	0.2127
	(0.346)	(0.154)	(0.346)	(0.154)
INTERN._COMP	0.2954	0.1536	0.2988	0.1601
	(0.232)	(0.104)	(0.231)	(0.104)
PATENTS_BOUGHT	-0.0083	0.1425	0.1006	0.1695
	(0.575)	(0.308)	(0.544)	(0.305)
PATENTS_SOLD	-1.7264**	-0.6184*	-1.9318**	-0.5793*
	(0.784)	(0.347)	(0.867)	(0.351)
INTERN._AGR.	0.3411	0.2850*	0.3510	0.2722
	(0.336)	(0.167)	(0.332)	(0.168)
PUBLIC_INCEN.	0.5513**	0.2565***	0.5535**	0.2561***
	(0.235)	(0.091)	(0.235)	(0.091)
LISTED	2.4816***	1.5039***	2.7548***	1.4223***
	(0.908)	(0.367)	(0.832)	(0.351)
BANK_SHARE	0.0113***	0.0031**	-0.0051	-0.0008
	(0.004)	(0.002)	(0.006)	(0.003)
NUM_BANKS	0.0482	0.0249	0.0520	0.0268
	(0.036)	(0.016)	(0.040)	(0.019)
LENGTH	-0.0136	-0.0004	0.0165	0.0014
	(0.011)	(0.004)	(0.012)	(0.006)
STRUMENTI_FIN	-0.5045	0.3072	1.2439*	0.5693*
	(0.551)	(0.252)	(0.731)	(0.336)
LOGSIZE	0.1747	0.2080***	0.0875	0.2174**
	(0.131)	(0.062)	(0.181)	(0.085)
HIGH_TECH	0.2062	0.2117**	0.2208	0.2192**
	(0.250)	(0.097)	(0.250)	(0.098)
NUM_BANKSxSMALL			-0.0079	-0.0016
			(0.049)	(0.021)
BANK_SHARExSMALL			0.0178***	0.0042
			(0.007)	(0.003)
LENGTHxSMALL			-0.0356**	-0.0026
			(0.016)	(0.007)
FIN._INSTRxSMALL			-2.0708**	-0.3200
			(0.992)	(0.434)
rationed$_{t-1}$		0.3238*		0.3423*
		(0.180)		(0.184)
Constant	-10.3087***	-4.6389***	-10.5396***	-4.5023***
	(2.050)	(0.796)	(2.021)	(0.808)
ρ	0.9414***	(0.0193)	0.9419***	(0.0192)
σ	2.5096***	(0.1929)	2.4956***	(0.1918)
ll	-35136.9		-35059.69	
N	584	1258	584	1258

Table 9: Estimation results: HECKMAN CONSIDERING EXTERNAL FINANCIAL NEED

	Intensity Eq(2)	**Propensity** Eq(1)	**Intensity** Eq(2)	**Propensity** Eq(1)
IE	0.2315***	0.1394***	0.2329***	0.1414***
	(0.031)	(0.031)	(0.031)	(0.031)
INVEST	-0.0028	0.0137**	-0.0011	0.0164***
	(0.011)	(0.006)	(0.011)	(0.006)
YOUNG	0.0589	0.0349	0.0153	0.0092
	(0.623)	(0.236)	(0.620)	(0.237)
M&As	0.6694**	0.1899	0.7073**	0.1828
	(0.339)	(0.151)	(0.338)	(0.146)
INTERN._COMP	0.2819	0.1556	0.2742	0.1523
	(0.232)	(0.104)	(0.232)	(0.104)
PATENTS_BOUGHT	-0.0242	0.1090	-0.0681	0.1279
	(0.557)	(0.308)	(0.554)	(0.308)
PATENTS_SOLD	-1.7275**	-0.6081*	-1.6896**	-0.6061*
	(0.776)	(0.343)	(0.781)	(0.346)
INTERN._AGR.	0.3489	0.2826*	0.3659	0.2738*
	(0.340)	(0.167)	(0.337)	(0.164)
PUBLIC_INCEN.	0.5459**	0.2431***	0.5455**	0.2429***
	(0.232)	(0.090)	(0.232)	(0.090)
LISTED	2.2482**	1.4324***	2.2022**	1.3657***
	(0.900)	(0.366)	(0.890)	(0.385)
LOGSIZE	0.1524	0.1965***	0.1306	0.1935***
	(0.131)	(0.062)	(0.130)	(0.062)
EFN	-4.3803**	-0.9732	-0.0462	3.0891
	(1.862)	(0.786)	(4.523)	(2.054)
BANK_SHARE	0.0115***	0.0031**	0.0070	-0.0023
	(0.004)	(0.002)	(0.029)	(0.013)
NUM_BANKS	0.0525	0.0260	0.2629	0.2777**
	(0.036)	(0.016)	(0.225)	(0.121)
LENGTH	-0.0132	-0.0001	0.0671	0.0636**
	(0.011)	(0.004)	(0.087)	(0.032)
FINANCIAL_INSTR	-0.4046	0.3407	-0.3583	0.3599
	(0.543)	(0.251)	(0.538)	(0.245)
HIGH_TECH	-0.0297	0.1643	-0.0374	0.1775*
	(0.261)	(0.105)	(0.261)	(0.105)
Constant	-7.6269***	-3.9733***	-9.6406***	-5.9070***
	(2.289)	(0.903)	(3.105)	(1.299)
NUM_BANKSxEFN			-0.4255	-0.4942**
			(0.436)	(0.235)
BANK_SHARExEFN			0.0094	0.0110
			(0.058)	(0.026)
LENGTHxEFN			-0.1608	-0.1264**
			(0.172)	(0.064)
g_sales$_{[t-1]}$		0.0490		0.0510
		(0.082)		(0.083)
ρ	0.9425***	(0.0186)	0.9449***	(0.0176)
σ	2.5057***	(0.1902)	2.5071***	0.2131
ll	-35113.2		-35011.31	
N	584	1258	584	1258

*p<0.10,** p<0.05,*** p<0.01
Note: Regressions include dummies for area

Table 10: Estimation results: HECKMAN CONSIDERING BANK COMPETITION

	Eq(2)	Eq(1)	Eq(2)	Eq(1)	Eq(2)	Eq(1)
IE	0.2369***	0.1397***	0.2367***	0.1406***	0.2317***	0.1402***
	(0.031)	(0.030)	(0.031)	(0.030)	(0.031)	(0.031)
INVEST	-0.0057	0.0125**	-0.0042	0.0123**	0.0008	0.0145**
	(0.010)	(0.006)	(0.010)	(0.006)	(0.010)	(0.006)
YOUNG	0.0124	0.0388	0.0177	0.0273	0.0353	0.0227
	(0.620)	(0.234)	(0.621)	(0.234)	(0.620)	(0.236)
M&As	0.6116*	0.1938	0.6218*	0.2181	0.6844**	0.2210
	(0.350)	(0.153)	(0.352)	(0.155)	(0.345)	(0.154)
INT_COMP	0.3119	0.1608	0.2968	0.1650	0.2854	0.1680
	(0.232)	(0.103)	(0.234)	(0.103)	(0.233)	(0.104)
PAT_B.	-0.0488	0.1275	0.0211	0.1598	0.0204	0.1536
	(0.580)	(0.304)	(0.590)	(0.307)	(0.564)	(0.312)
PAT_S.	-1.7752**	-0.6272*	-1.7274**	-0.6272*	-1.7577**	-0.6414*
	(0.786)	(0.344)	(0.814)	(0.348)	(0.799)	(0.348)
INT_AGR	0.3317	0.2848*	0.3251	0.2853	0.3143	0.2698
	(0.341)	(0.167)	(0.343)	(0.175)	(0.348)	(0.172)
PUB_INC	0.5360**	0.2455***	0.5788**	0.2636***	0.5397**	0.2467***
	(0.238)	(0.091)	(0.238)	(0.091)	(0.236)	(0.091)
LISTED	2.5393***	1.4934***	2.5278***	1.5434***	2.3140**	1.5251***
	(0.878)	(0.360)	(0.932)	(0.383)	(0.909)	(0.383)
LOGSIZE	0.1866	0.2012***	0.1823	0.1964***	0.1628	0.1986***
	(0.131)	(0.062)	(0.133)	(0.062)	(0.131)	(0.062)

It continues next page

Table 10 continues

	Eq(2)	Eq(1)	Eq(2)	Eq(1)	Eq(2)	Eq(1)
BANK_SH	0.0111***	0.0030*	0.0114***	0.0031**	0.0114***	0.0031**
	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
NUM_BAN	0.0462	0.0234	0.0509	0.0247	0.0510	0.0258
	(0.037)	(0.016)	(0.037)	(0.016)	(0.036)	(0.016)
LENGTH	-0.0134	-0.0002	-0.0141	0.0002	-0.0130	0.0005
	(0.011)	(0.004)	(0.011)	(0.004)	(0.011)	(0.004)
FIN_INS.	-0.5095	0.3116	-0.5148	0.3281	-0.4979	0.3350
	(0.544)	(0.247)	(0.548)	(0.247)	(0.534)	(0.249)
HT	0.2105	0.2175**	0.2009	0.2147**	-0.0138	0.1740*
	(0.251)	(0.097)	(0.252)	(0.097)	(0.262)	(0.105)
Cons.	-10.3506***	-4.5665***	-10.2864***	-4.2505***	-15.8254***	-8.0977***
	(2.003)	(0.791)	(2.177)	(0.850)	(5.566)	(2.141)
BANK_CO	-39.4754	-6.0762				
	(41.102)	(17.496)				
CH3			-0.4454	-0.8572*	14.6859	7.1774**
			(1.223)	(0.501)	(9.126)	(3.533)
EFN					11.8780	7.5562**
					(9.705)	(3.726)
CH3xEFN					-29.7625*	-15.6612**
					(17.688)	(6.781)
g_sales$_{[t-1]}$		0.0447		0.0345		0.0368
		(0.083)		(0.078)		(0.078)
ρ	0.9447***	(0.176)	0.9487***	(0.0173)	0.9463***	(0.0180)
σ	2.526***	(0.1911)	2.5391***	(0.1930)	2.5121***	(0.1903)
ll	-35156.65		-35109.31		-35002.96	
N	584	1258	584	1258	584	1258

Table 11: Estimation results: CONDITIONAL LOGIT

In this model only switchers - that is, firms that introduced an innovation in just one of the two periods - contribute to the likelihood function. It controls for unobserved firm characteristics that are constant over time. The dependent variable is a dummy variable equal to 1 if the firm introduced an innovation in just one of the two periods

	(1)	(2)
IE	0.0942	
	(0.091)	
INVEST	0.0592***	
	(0.020)	
R&D		1.1852***
		(0.295)
INVE		0.0705***
		(0.022)
PATENTS_SOLD	-1.4270	-0.9091
	(1.185)	(1.031)
PATENTS_BOUGHT	1.0923	0.4738
	(1.671)	(1.057)
INTERN._AGR.	1.6552***	1.5007**
	(0.534)	(0.612)
M&As	1.1833***	0.7833*
	(0.385)	(0.436)
INTERN._COMP	0.3529	0.3976
	(0.393)	(0.399)
LOGSIZE	0.5377	0.6246
	(0.541)	(0.543)
PUBLIC_INCEN.	0.4973**	0.5279**
	(0.238)	(0.248)
LISTED	2.6798**	2.5719**
	(1.351)	(1.301)
BANK_SHARE	0.0448**	0.0430**
	(0.021)	(0.018)
NUM_BANKS	0.1362	0.0741
	(0.107)	(0.110)
LENGTH	-0.0074	0.0003
	(0.019)	(0.002)
FIN._INSTR	0.6966	0.9219*
	(0.487)	(0.501)
BANK_SHARE*SMALL	-0.0478**	-0.0481***
	(0.021)	(0.018)
NUM_BANKS*SMALL	-0.0129	0.0748
	(0.115)	(0.118)
LENGTH*SMALL	0.0116	0.0028
	(0.020)	(0.002)
SALES	0.0000	-0.0000
	(0.000)	(0.000)
ll	-4161.4137	-4211.8384
N	320	344

*p<0.10,** p<0.05, ***p<0.01

Table 12: Estimation results: CONDITIONAL LOGIT
In this model only switchers - that is, firms that introduced an innovation in just one of the two periods - contribute to the likelihood function. It controls for unobserved firm characteristics that are constant over time. The dependent variable is a dummy variable equal to 1 if the firm introduced an innovation in just one of the two periods

	(1)	(2)
IE	0.0900	
	(0.101)	
INVEST	0.0589***	
	(0.021)	
R&D		1.1740***
		(0.292)
INVE		0.0676***
		(0.022)
PATENTS_SOLD	-1.3398	-0.8299
	(1.310)	(1.152)
PATENTS_BOUGHT	1.7168	0.6122
	(1.677)	(0.958)
INTERN._AGR.	1.7995***	1.6157***
	(0.553)	(0.590)
M&As	1.1113***	0.6764
	(0.402)	(0.448)
INTERN._COMP	0.4203	0.4472
	(0.376)	(0.373)
LOGSIZE	0.8827**	0.9098**
	(0.442)	(0.459)
PUBLIC_INCEN.	0.5625**	0.6107**
	(0.239)	(0.249)
LISTED	0.9050	1.0268
	(1.179)	(1.114)
BANK_SHARE	-0.0064	-0.0103*
	(0.006)	(0.006)
NUM_BANKS	0.1015	0.1081
	(0.091)	(0.105)
LENGTH	0.0034**	0.0029***
	(0.002)	(0.001)
FIN._INSTR	0.6187	0.8385
	(0.507)	(0.536)
BANK_SHARE*HIGH_TECH	0.0142	0.0223**
	(0.013)	(0.011)
NUM_BANKS*HIGH_TECH	0.2271	0.1813
	(0.209)	(0.230)
LENGTH*HIGH_TECH	-0.0031	-0.0025**
	(0.002)	(0.001)
SALES	-0.0000	-0.0000
	(0.000)	(0.000)
ll	-4191.3611	-4124.6814
N	320	344

*p<0.10,** p<0.05,*** p<0.01

Unless otherwise expressly stated, all original material of whatever nature created by Caterina Giannetti and included in this thesis, is licensed under a Creative Commons Attribution Noncommercial Share Alike 2.5 Italy License.

Check creativecommons.org/licenses/by-nc-sa/2.5/it/ for the legal code of the full license.

Ask the author about other uses.