Imperfect Competition and Market Structure with Asymmetric Information: the Italian Banking Sector

by

Nicola Pavanini

Thesis

Submitted to the University of Warwick

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Economcis

June 2014

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

There is a long list of people who I am grateful to for the completion of this PhD thesis. Foremost, I have been very lucky to receive the guidance of my main advisor, Greg Crawford. I am highly indebted to Greg, who gave me great support, advice, and critics on my work, but most importantly has put in a sincere effort to grow me into a valid researcher. I also want to thank my other formal and informal advisors, Fabiano Schivardi, Carlos Noton, and Mike Waterson, for their invaluable help throughout this long journey.

I express my gratitude to the Warwick Economics community, which provided me with an excellent academic environment. I learned a lot thanks to Professors like Rocco Macchiavello and Chris Woordruff, and from PhD fellows and friends as Alessandro, Federica, Guillaume, Kazim, and Tasos. I acknowledge the financial support of the CAGE research centre at Warwick Economics Department.

I feel also deeply indebted to my parents, for providing me with such a great life example, and for always supporting and encouraging me throughout all these years. They deserve all the credit for what I've been able to do. Last but not least, I have been extremely lucky to cross my path with my girlfriend June, who brought happiness and love to my life.

Declarations

I declare that the thesis has not been submitted for a degree at any other university, and that it contains work based on collaborative research.

Chapter 2 *Asymmetric Information and Imperfect Competition in the Loan Market* was written with Prof. Gregory S. Crawford (University of Zürich) and with Prof. Fabiano Schivardi (LUISS and EIEF). This has been published in September 2013 in the Working Paper Series of the Centre for Competitive Advantage in the Global Economy, Department of Economics, University of Warwick, paper number 167/2013.

Chapter 3 *Dynamic Entry and Exit with Learning by Branching* was written with Prof. Fabiano Schivardi (LUISS and EIEF).

Chapter 4 *Market Structure and Multi Market Contact* was written with Dr. Lorenzo Ciari (EBRD, London).

Abstract

This thesis studies the relationship between asymmetric information, imperfect competition, and market structure in the Italian banking sector. In three coherently connected chapters it extends and adapts different structural models from the literature in empirical industrial organization to the special case of the credit market, introducing informational asymmetries between lenders and between borrowers and lenders.

The first chapter gives an introductory overview of the thesis, outlying the fundamental contribution of each paper.

In the second chapter, we measure the consequences of asymmetric information in the Italian market for small business lines of credit. We estimate models of demand for credit, loan pricing, loan use, and firm default based on the seminal work of Stiglitz and Weiss [1981]. Preliminary results suggest evidence of asymmetric information, separately identifying adverse selection and moral hazard. We use our results to quantify the impact of asymmetric information on pricing and welfare, and the role imperfect competition plays in mediating these effects.

In the third chapter, we look at whether asymmetric information is a potential determinant of market structure in the banking industry. We measure welfare under different counterfactual scenarios, with and without borrower-lender asymmetric information and reducing incumbents' informational advantage. We develop a dynamic structural game of banks' entry, exit and investment with learning by branching based on Weintraub et al. [2008b], together with a static framework of firms' demand for credit, loan size, default and banks' pricing, that allows us to identify the effect of asymmetric information on market structure.

In the fourth chapter, we measure the impact of endogenous multi market contact on entry decisions of Italian national banks. We develop a static model of market structure with incomplete information as in Seim [2006], allowing for global players' heterogeneity and spatial correlation of entry decisions across different local markets. Preliminary results show that multi market contact enhances banks' profitability, suggesting that it might facilitate implicit collusion as in Bernheim and Whinston [1990].

The fifth chapter concludes, summarizing the main findings and tracing out the directions for future research.

Chapter 1

Introduction

Asymmetric information is an important matter in insurance and credit markets. Theoretical work on asymmetric information has grown extensively since the 1970s, starting from the pioneering papers by Arrow [1963], Akerlof [1970] and Rothschild and Stiglitz [1976], explaining the inefficiencies caused by this source of market failure. For the credit market in particular, Stiglitz and Weiss [1981] have shown how asymmetric information can lead to an equilibrium with credit rationing. However, the empirical counterpart has gained increasing interest only in the last decade. This delay is partially because of lack of data, but mostly because asymmetric information is by definition something that is hard to measure. In fact, adverse selection is based on hidden information, and moral hazard is based on hidden actions. So far, the empirical literature has focussed mainly on testing for selection in insurance markets, and on estimating the welfare consequences of detected selection and of potential public policy interventions.[1]

One of the first important empirical contributions is the positive correlation test for asymmetric information, developed by Chiappori and Salanié [2000]. This test compares claim rates for consumers who self select into different insurance contracts. There is evidence of either adverse selection or moral hazard if, conditional on all the observable characteristics of consumers, those who select more coverage have also higher claim rates. This is identified through the correlation between the unobservables that determine demand for coverage and the unobservables that determine claim behavior. However, this method cannot disentangle the two forms of asymmetric information, and doesn't provide the necessary structure to determine the efficiency of the market through counterfactual policy experiments. For this reason, the literature has developed a more structural approach that incorporates

[1] Einav et al. [2010a] provide an excellent survey of empirical models of insurance markets.

heterogeneity in consumer preferences and risk, allowing to do welfare analysis.

Also these structural frameworks are based on the correlation between unobservables in models of consumers' choice of insurance and their following claim behavior. However, the use of richer theoretical foundations allows to answer to a wider range of questions. One area of interest focuses on the expected utility theory, separately identifying heterogeneity in risk and risk aversion (Cohen and Einav [2007], Einav et al. [2010b]). Another area relies on discrete choice models, usually applied in empirical industrial organization, and tries to combine asymmetric information and imperfect competition (Lustig [2011], Starc [2013], Einav et al. [2012]).

The second chapter of this thesis contributes to the latter approach, investigating the welfare consequences of asymmetric information in the Italian market for small business lines of credit. It extends the model of Einav et al. [2012] to the Italian oligopolistic banking sector, showing how market power can mitigate the inefficiencies of asymmetric information between borrowers and lenders. It develops a structural model of firms' demand for credit, loan use and default and of banks' pricing that separately identifies adverse selection and moral hazard.

The third chapter of this thesis builds on the framework of the second chapter to analyze the effects of asymmetric information on market structure in the banking industry. It combines the theoretical (Dasgupta and Stiglitz [1988], Cabral and Riordan [1994], Besanko et al. [2010]) and empirical (Benkard [2000], Benkard [2004]) literature on the strategic implication of learning by doing into a dynamic structural game of banks' entry, exit and investment with learning by branching based on Weintraub et al. [2008b] and Weintraub et al. [2008a], which allows incumbent banks to learn about their borrowers and acquire an informational advantage over potential entrants. This can in turn lead to informational entry barriers and concentration in the banking sector.

The fourth chapter investigates the effect of multi market contact between the main Italian credit institutions on market structure. Since the theoretical contribution of Bernheim and Whinston [1990], who formalized the collusive gains from multiple contacts across markets, most of the empirical literature has focussed on the impact of multiple links on firms' profits, taking market structure as given. This work instead endogenizes market structure and the network of contacts among players. It allows banks to have private information about their profitability, as in Seim [2006], and to decide their branching strategy based on the links with their rivals in neighboring markets.

The last chapters draws the conclusions of this thesis and outlines the directions for future work.

Chapter 2

Asymmetric Information and Imperfect Competition in Lending Markets

Asymmetric Information and Imperfect Competition in Lending Markets*

Gregory S. Crawford†, Nicola Pavanini‡, Fabiano Schivardi§

June 2014

*We thank Daniel Ackerberg, Pierre-André Chiappori, Lorenzo Ciari, Valentino Dardanoni, Ramiro de Elejalde, Liran Einav, Moshe Kim, Rocco Macchiavello, Carlos Noton, Tommaso Oliviero, Steven Ongena, Ariel Pakes, Andrea Pozzi, Pasquale Schiraldi, Matt Shum, Michael Waterson, Chris Woodruff, Christine Zulehner and seminar participants at Warwick, PEDL, Barcelona GSE Banking Summer School 2012, EARIE 2012, EUI, Tilburg, Zürich, Bocconi, 2014 Winter Marketing-Economics Summit in Wengen, IO session of the German Economic Association in Hamburg, and St. Gallen for helpful comments. We thank for financial support the Research Centre Competitive Advantage in the Global Economy (CAGE), based in the Economics Department at University of Warwick.

†University of Zürich, CEPR and CAGE, gregory.crawford@econ.uzh.ch

‡University of Zürich, nicola.pavanini@econ.uzh.ch

§LUISS, EIEF and CEPR, fschivardi@luiss.it

Abstract

We measure the consequences of asymmetric information and imperfect competition in the Italian market for small business lines of credit. We provide evidence that a bank's optimal price response to an increase in adverse selection varies depending on the degree of competition in its local market. More adverse selection causes prices to increase in competitive markets, but can have the opposite effect in more concentrated ones, where banks trade off higher markups and the desire attract safer borrowers. This implies both that imperfect competition can moderate the welfare losses from adverse selection, and that adverse selection can moderate the welfare losses from market power. Exploiting detailed data on a representative sample of Italian firms, the population of medium and large Italian banks, individual lines of credit between them, and subsequent defaults, we estimate models of demand for credit, loan pricing, loan use, and firm default to measure the extent and consequences of asymmetric information in this market. While our data include a measure of observable credit risk available to a bank during the application process, we allow firms to have private information about the underlying riskiness of their project. This riskiness influences banks' pricing of loans as higher interest rates attract a riskier pool of borrowers, increasing aggregate default probabilities. We find evidence of adverse selection in the data, and conduct a policy experiment to double its magnitude. As predicted, in this counterfactual scenario equilibrium prices rise in more competitive markets and decline in more concentrated ones.

2.1 Introduction

Following the seminal work of Akerlof [1970] and Rothschild and Stiglitz [1976], a large theoretical literature has stressed the key role of asymmetric information in financial markets. This literature has shown that asymmetric information can generate market failures such as credit rationing, inefficient provision, misprincing of risk and, in the limit, market breakdown.[1] Indeed, the recent financial crisis can be seen as an extreme manifestation of the problems that asymmetric information can cause. Deepening our understanding of the extent and causes of asymmetric information is key for the design of a regulatory framework that limits their negative consequences.

Although the basic theoretical issues are well understood, empirical work is fairly rare. Asymmetric information is by definition hard to measure. If a financial intermediary, such as an insurer or a lender, has an information disadvantage with respect to a potential insuree/borrower, it is very unlikely that such a disadvantage can be overcome by the researcher, if not in experimental settings. While one cannot generally construct measures of the ex-ante unobserved characteristics determining riskyness, it is often possible to observe ex-post outcomes, such as filing a claim to an insurance company or defaulting on a loan. The empirical literature has been built on these facts, analyzing how agents with different ex-post outcomes self select ex-ante into contracts (if any) with different characteristics in terms of price, coverage, deductibles etc. (Chiappori and Salanié [2000], Abbring et al. [2003], Lustig [2011], Einav et al. [2012], Starc [2013]).

We measure the consequences of asymmetric information and imperfect competition in the Italian market for small business lines of credit. We exploit detailed, proprietary data on a representative sample of Italian firms, the population of medium and large Italian banks, individual lines of credit between them, and subsequent individual defaults. While our data include a measure of observable credit risk comparable to that available to a bank during the application process, in our model we allow firms to have private information about the underlying riskiness of the project they seek to finance. The market is characterized by adverse selection if riskier firms are more likely to demand credit. As shown by Stiglitz and Weiss [1981], in this setting an increase in the interest rate exacerbates adverse selection, inducing a deterioration in the quality of the pool of borrowers. We formulate and structurally estimate a model of credit demand, loan size, default, and bank pricing based on the insights in Stiglitz and Weiss [1981] that allows us to estimate the ex-

[1] See, for example, Banerjee and Newman [1993], Bernanke and Gertler [1990], DeMeza and Webb [1987], Gale [1990], Hubbard [1998], Mankiw [1986], Mookherjee and Ray [2002].

tent of adverse selection in the market and to run counterfactuals that approximate economic environments of likely concern to policymakers.

One key contribution of our paper is that we study adverse selection in an imperfectly competitive market. This differs from most of the previous literature, that, due to data limitation or to specific market features, has assumed either perfectly competitive markets, or imperfectly competitive markets subject to significant regulatory oversight. Assuming perfect competition in the market for small business loans is not desirable, given the local nature of small business lending and the high degree of market concentration at the local level, the latter due to entry barriers in the Italian banking sectors that persisted into the 1990s. We show that the degree of competition can have significant consequences on the equilibrium effects of asymmetric information. Intuitively, with perfect competition banks price at average costs (e.g. Einav and Finkelstein (2011)). When adverse selection increases, the price also rises, as a riskier pool of borrowers implies higher average costs in the form of more defaults. When banks exert market power, however, greater adverse selection can *lower* prices, as it implies a riskier pool of borrowers at any given price, lowering infra marginal benefits of price increases in the standard (e.g. monopoly) pricing calculus. This implies both that imperfect competition can moderate the welfare losses from asymmetric information and that adverse selection can moderate the welfare losses of market power.

To analyze these questions, we construct a model where banks offer standardized contracts to observationally equivalent firms. Loan contracts are differentiated products in terms of, among other characteristics, the amount granted, a bank's network of branches, the years a bank has been in a market, and distance from the closest branch. Banks set interest rates by competing Bertrand-Nash. Firms seek lines of credit to finance the ongoing activities associated with a particular business project, the riskiness of which is private information to the firm. Firms choose the preferred loan, if any, according to a mixed logit demand system. They also choose how much of the credit line to use. Finally, they decide if to repay the loan or default. The degree of adverse selection is determined by two correlations: that between the unobservable determinants of the choice to take up a loan and default and that between unobserved determinants of how much of that loan to use and default. For a given interest rate, firms' expected profits are increasing with risk due to the insurance effect of loans: banks share a portion of the costs of unsuccessful projects. As a result, higher-risk firms are more willing to demand higher-rate loans. This, in turn, influences the profitability of rate increases by banks.[2] We show with a Monte

[2]Handel [2013], Lustig [2011], and Starc [2013] find similar effects of adverse selection and

Carlo simulation that imperfect competition can indeed mitigate the effects of adverse selection.[3] The effects of asymmetric information on prices depends on market power. When markets are competitive, more asymmetric information always leads to higher rates and less credit. As banks' market power increases, this relationship becomes weaker and eventually turns negative.

We estimate the model on highly detailed microdata from the Bank of Italy covering individual loans between firms and banks between 1988 and 1998. There are two key elements of this data. The first, from the Italian Central Credit Register (*Centrale dei Rischi*), provides detailed information on all individual loans extended by the 90 largest Italian banks (which account for 80% of the loan market), including the identity of the borrower and interest rate charged. It also reports whether the firm subsequently defaulted. The second, from the *Centrale dei Bilanci* database, provides detailed information on borrowers' balance sheets. Critically, this second dataset includes an observable measure of each firm's default risk (SCORE). Combining them yields a matched panel dataset of borrowers and lenders. While the data span a 11-year period and most firms in the data take out multiple loans, in our empirical analysis, we only use the first year of each firm's main line of credit. This avoids the need to model the dynamics of firm-bank relationships and the inferences available to subsequent lenders of existing lines of credit.[4] We define local markets at the level of provinces, administrative units roughly comparable to a US county that, as discussed in detail by Guiso et al. [2013], constitute a natural geographical unit for small business lending. We estimate individual firms' demand for credit, banks' pricing of these lines, firm's loan use and subsequent default. We extend the econometric approach taken by Einav et al. [2012] to the case of multiple lenders by assuming unobserved tastes for credit independent of the specific bank chosen to supply that credit.

and the literature on demand estimation for differentiated products Berry [1994]; Berry et al. [1995]; Goolsbee and Petrin [2004]. Data on default, loan use, demand, and pricing separately identify the distribution of private riskiness from heterogeneous firm disutility from paying interest.

We find that the choice to borrow, the amount used and the decision to

imperfect competition in US health insurance markets. Each of these focuses on the price-reducing effect of asymmetric information in the presence of imperfect competition. None articulates the non-monotonicity of these effects depending on the strength of competition, an empirically relevant result in our application.

[3] In the Monte Carlo we vary the degree of competition changing the number of banks in the market, as well as varying the price sensitivity of borrowers, which increases/decreases their utility from the outside option of not borrowing.

[4] A similar approach is followed, among others, by Chiappori and Salanié [2000]. We model the dynamics of firm-bank relationships in a companion paper Pavanini and Schivardi [2013].

default depend on observables as expected. In particular, a higher interest rate reduces the probability that a firm borrows but, conditional on borrowing, increases the default probability. Among other observables, older firms are both less likely to demand credit, arguably because they have more internally generated funds, and more likely to default. Firms with larger assets demand more credit and default less. In terms of correlation in unobservables, we find a positive correlation between the choice to borrow and default, and between how much loan to use and default. We simulate with a counterfactual experiment the possible consequences of a credit crunch, where risky firms become more exposed to financial distress than safe ones and demand more credit. Our results show that when we increase the correlation in the unobservables (thus increasing the extent of adverse selection), prices in most markets increase, but they fall in some markets. The change in prices is related to different measures of market concentration,[5] supporting the view that market concentration can mitigate the negative effects of asymmetric information. As a consequence of this price decrease, the share of borrowing firms in more concentrated markets increases, and their average default rate falls.

This paper contributes to two main strands of empirical work. The first is the literature on empirical models of asymmetric information, so far mainly focussed on insurance markets. We look at the less developed area of credit markets, where the most recent applications have followed both experimental (Karlan and Zinman [2009]) and structural (Einav et al. [2012]) approaches. Our novelty is to introduce imperfect competition. We show that this is important, as the impact of asymmetric information depends crucially on the nature of competition in the market. The second field we contribute to is the literature on empirical banking, where we are not aware of any structural model that seeks to measure the consequences of asymmetric information and the role competition plays in mediating its effects. Nonetheless, several reduced form papers on Italian banking provide motivation for a model that structurally combines these two effects. For example, Bofondi and Gobbi [2006] show evidence that new banks entering local markets perform poorly relative to incumbents, as entrants experience higher default rates and concentration and default rates are positively correlated. Gobbi and Lotti [2004] claim that there is a positive correlation between branching and markets with low proprietary information services, and that interest rate spreads are positively related to entry of de novo banks, but not of banks existing in other markets. Finally, Panetta et al. [2009] show that mergers enhance pricing of observable risk, as merged banks achieve a better match

[5] In the counterfactuals we relate the equilibrium price variation to the estimated markups from the demand model. We also experiment with HHI in terms of branches and loans.

of interest rates and default risk, mainly due to better information processing.

The structure of the paper is the following. In Section 2 we describe the dataset and the market, in Section 3 we present the reduced form tests of adverse selection, Section 4 outlines the structural model, and Section 5 describes the econometric specification of demand, loan size, default and supply. The estimation and the results are in Section 6, the counterfactuals are in Section 7, Section 8 concludes.

2.2 Data and Institutional Details

We use a unique dataset of small business credit lines, previously used in Panetta et al. [2009].[6] It is based on three main sources of data. Interest rate data and data on outstanding loans are from the Italian *Centrale dei Rischi*, or Central Credit Register. Firm-level balance sheet data are from the *Centrale dei Bilanci* database. Banks' balance-sheet and income-statement data are from the Banking Supervision Register at the Bank of Italy. By combining these data, we obtain a matched panel dataset of borrowers and lenders extending over an eleven-year period, between 1988 and 1998. We also collected data on bank branches at the local level since 1959.

The Central Credit Register (hereafter CR) is a database that contains detailed information on individual bank loans extended by Italian banks. Banks must report data at the individual borrower level on the amount granted and effectively utilized for all loans exceeding a given threshold,[7] with a breakdown by type of the loan (credit lines, financial and commercial paper, collateralized loans, medium and long-term loans and personal guarantees). Banks also report if they classify a loan as bad, meaning that they attach a low probability to the event that the firm will be able to repay the loan in full. We define a default as a loan being classified as bad.[8] In addition, a subgroup of around 90 banks (accounting for more than 80 percent of total bank lending) have agreed to file detailed information on the interest rates they charge to individual borrowers on each type of loan.

We restrict our attention to short-term credit lines, which have ideal features for our analysis. First, the bank can change the interest rate at any time, while the borrower can close the credit line without notice. This means that differences between the interest rates on loans are not influenced by differences in the maturity of the loan. Second, the loan contracts included in the CR are homogeneous products, so that they can be meaningfully compared across banks and firms. Third, they are not collateralized, a key feature for our analysis, as adverse selection issues become less relevant for collateralized borrowing. Fourth, short term bank loans are the main source of borrowing of Italian firms. For example, in 1994 they represented 53 percent of the total debts according to the Flow of Funds data. We define the interest rate as the ratio of the payment made in each year by the firm to the bank

[6] For reasons that will be explained below, in this paper we only use on a subset of the original data. This section focusses on the description of this subset, referring the interested reader to Panetta et al. [2009] for descriptive statistics of the full dataset.

[7] The threshold was 41,000 euros (U.S. $42,000) until December 1995 and 75,000 euros thereafter.

[8] We do not observe if a loan actually reverts to not being bad. However, this seems to be a rather unlikely event. Moreover, classifying a loan as bad has a negative impact on bank accounting ratios, even before the firm formally defaults. So this is clearly a costly event in itself for the bank.

to the average amount of the loan. The interest payment includes the fixed expenses charged by the bank to the firm (e.g. which encompass the cost of opening the credit line or the cost of mailing the loan statement).

We focus on a subsample of the available data, namely on the main credit line of the first year a firm opens at least one credit line. Considering only the first year is a common assumption in static empirical models of insurance with asymmetric information, starting from Chiappori and Salanié [2000]. This is done to avoid modeling heterogenous experience ratings among borrowers and loan renegotiation, challenging topics, and ones that we leave for future research. Moreover, we focus on the main new credit line because it accounts on average for around 75% of the total share of new yearly credit (both usable and used),[9] even if in Italy multiple relationship banking is widely used by firms to reduce liquidity risk (Detragiache et al. [2000]). This means that we restrict our attention only to the first year in which we observe a firm in our data.[10] This reduces the sample size from around 90,000 firms to over 40,000.[11] Table 2.1, Panel A reports the loan level information that we use in the empirical analysis. Out of over 20,000 potential borrowers, 36% take up a loan in our sample period, and use on average 80% of the amount granted. Of these, around 15% end up being classified as bad loans within the following 3 years.[12] The average amount granted is 350,000 euros, and the average interest rate charged is just below 15%.

Panel B of Table 2.1 shows summary statistics for the 90 reporting banks. The average total assets level is almost 11 billions, they employ 3,200 employees and have a share of bad loans over total loans of 6%. The average bank is present in 34 provinces out of 95, but with great variation across banks.

The *Centrale dei Bilanci* (hereafter CB) collects yearly data on the balance sheets and income statements of a sample of about 35,000 Italian non-financial and non-agricultural firms. This information is collected and standardized by the CB, that sells these data to banks for their lending decisions. The unique feature of the CB data set is that, unlike other widely used data sets on individual companies (such as the Compustat database of US companies), it has wide coverage of small

[9] The main line is defined as the line for which the amount used, regardless of the amount granted, is the highest. For cases in which multiple lines have the same amount used, then the one with the lowest price is chosen.

[10] To avoid left censoring issues we drop the first year of our sample (1988) and just look at new relationships starting from 1989.

[11] Due to computational constraints, we are able to estimate the model in this version of the paper only on half of the sample. Therefore we randomly pick 50% of the province-year combinations in our sample.

[12] We classify a borrower as defaulter when any of its loans is pass due within the next 3 years from the initial borrowing date.

and medium companies; moreover, almost all the companies in the CB sample are unlisted. The coverage of these small firms makes the data set particularly well suited for our analysis, because informational asymmetries are potentially strongest for these firms. Initially, data were collected by banks themselves and transmitted to the CB. In time, the CB has increased the sample size drawing from balance sheets deposited with the commerce chambers (limited liability companies are obliged to file their balance sheets to the commerce chambers, that make them available to the public). The database is fairly representative of the Italian non-financial sector. The firms in the CB sample represent about 49.4% of the total sales reported in the national accounting data for the Italian non-financial, non-agricultural sector. In addition to collecting the data, the CB computes an indicator of the risk profile of each firm (which we refer to in the remainder of this paper as the SCORE). The SCORE represents our measure of a firm's observable default risk. It takes values from 1 to 9 and is computed annually using discriminant analysis based on a series of balance sheet indicators (assets, rate of return, debts etc.) according to the methodology described in Altman [1968] and Altman et al. [1994].

We defined a borrowing firm as one that shows up as a borrower in the CR database. Non borrowing firms are defined according to two criteria: (a) they are not in the CR database; (b) they report zero bank borrowing in their balance sheets. We use the second definition to exclude firms that are not in our CR database but are still borrowing from banks, either from one of the non-reporting banks or through different loan contracts.[13] Table 2.1, Panel C reports descriptive statistics for the sample of borrowing and non-borrowing firms. These two groups of firms appear to be fairly similar, except that borrowing firms seems to have more fixed assets and be slightly younger on average. In terms of bank relations, our sample of borrowing firms have on average around 3.4 credit lines active every year. They open one new line every year and close 0.6. Note that these firms are mostly new borrowers, so they are more likely to be in the process of expanding their number of relationships. The share of credit used from the main line is around 70%, and it goes up to 75% when a firm borrows for the first year. This shows that focusing on the main line captures most of the credit that firms borrow, especially for new firms.

There is ample evidence that firms, particularly small businesses like the ones in our sample, are tied to their local credit markets. For instance, Petersen and Rajan [2002] and Degryse and Ongena [2005] show that lending to small businesses is a highly localized activity as proximity between borrowers and lenders facilitates

[13] This implies that we exclude from our sample around 27,000 firms that borrow from banks not included in our sample, or borrow from the banks in our sample but using a different type of loan. This might be a possible source of selection bias that we will need to investigate.

information acquisition. Segmentation of local credit markets is thus very likely to occur. In our market definition we will use provinces as our geographical units. Provinces are administrative unit roughly comparable to a US county. They are a proper measure of local markets in banking for at least three reasons. First, this was the definition of a local market used by the Bank of Italy to decide whether to authorize the opening of new branches when entry was regulated. Second, according to the Italian Antitrust authority the "relevant market" in banking for antitrust purposes is the province. Third, the bankers' rule of thumb is to avoid lending to a client located more than three miles from the branch. At the time of our data, there were 95 provinces. We report summary statistics of markets (defined more precisely below) in Panel D of Table 2.1, which shows that there are almost 6 banks per province-year in our sub-sample, each bank has on average almost 19 branches per province, with a market share of 7% for both branches and loans. On average a bank has been in a province for 22 years.[14]

[14] We start counting the years from 1959, which is the first year that we observe in the branching data.

Table 2.1: Summary Statistics

	Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Panel A:	Demand	20,080	0.36	0.48			
Loan Level	Loan Size	7,170	0.81	1.34			
	Default	7,170	0.15	0.36			
	Amount Granted	7,170	352.85	409.68			
	Interest Rate	7,170	14.67	4.03			
Panel B:	Total Assets	900	10,726.8	16,965.6			
Bank Level	Employees	896	3,179.9	4,582.5			
	Bad Loans	893	6.2	6.3			
	Number of Provinces	861	34.54	30.19			
Panel C:		**Borrowing Firms**			**Non-Borrowing Firms**		
Firm Level	Fixed Assets	7,170	2,631.52	11,136.64	12,876	1,597.84	7,705.57
	Intangible/Tot Assets	7,170	0.19	0.25	12,876	0.19	0.27
	Net Worth	7,170	1,441.18	4,683.39	12,876	1,591.21	6,845.24
	Trade Debit	7,170	1,402.73	4,197.48	12,876	1,358.50	6,723.18
	Profits	7,170	744.90	2,043.97	12,876	517.79	2,636.62
	Cash Flow	7,170	441.37	1,754.81	12,876	474.47	2,362.71
	Firm's Age	7,170	11.68	11.83	12,876	13.76	13.37
	Branch distance (km)	7,170	2.64	6.56			
	Number of Lenders	31,328	3.40	2.36			
	Lines Opened	31,328	1.04	1.55			
	Lines Closed	31,328	0.61	1.23			
	Share of Main Line	26,776	0.70	0.25			
	Share of Main New Line	6,095	0.74	0.25			
Panel D:	Number of Banks	386	5.90	4.11			
Market level	Number of Branches	2,279	18.77	31.34			
	Share of Branches	2,279	0.07	0.09			
	Years in Market	2,279	22.21	13.77			
	Market Shares	2,279	0.07	0.08			

Note: In Panel A an observation is a firm for the first variable and a loan contract for the others. Demand is a dummy for taking a loan or not, loan size is the share of amount used over granted, default is a dummy for a firm having any of its loans classified as bad at most within 3 years from demanding the loan we consider, amount granted is in thousands of euros. In Panel B an observation is a bank-year. Employees is the number of employees at the end of the year. Bad loans is a percentage of total loans. In Panel C an observation is a firm for the first 8 variables and a firm-year for the others. The balance sheet variables in this panel are winsorized at the 1st and 99th percentile. The SCORE is the indicator of the risk of the firm computed each year by the CB (higher values indicate riskier companies). Number of lenders is the number of banks from which the firm borrows through these credit lines. The last two variables represent the ratio of credit utilized from the main line over total credit utilized, when credit utilized is non-zero. In Panel D an observation is year-province for the number of banks, and bank-year-province for the other variables. Number and share of branches are per bank-province-year, years in market are the number of years a bank has been in a province for since 1959. Market shares are in terms of loans.

2.3 Reduced Form Evidence

We conduct some reduced form analysis to test for evidence of asymmetric information and to justify the use of a structural model. We follow the early empirical literature on positive correlation tests introduced by Chiappori and Salanié [2000]. We propose two tests, one based on the choice to take up a loan and another based on the choice of how much to draw on the credit line. Both tests are based on the correlation between the unobservables driving these choices and the unobservables influencing default. The choice of these tests gives a flavor of the identification strategy that we will rely on in the structural model, explained in Section 4. We run these tests on the whole sample and for the first loan ever taken, the set of loans that we will use in the structural estimates.

2.3.1 Demand and Default

We start by investigating whether firms that are more likely to demand credit are also more likely to default. The CB dataset includes both firms borrowing and not borrowing, while we only observe default on the loan only for borrowing firms. We can formalize the problem as a two equations selection model:

$$\begin{aligned} d_i &= \mathbf{1}(X_i^d\beta + \nu_i > 0) \\ f_i &= \mathbf{1}(X_i^f\gamma + \eta_i > 0) \end{aligned} \tag{2.1}$$

where d_i is equal to 1 if the firm borrows and f_i is equal to one if the borrower is a defaulter [15] and is observed only if $d_i = 1$. This is similar to the classical selection model analyzed by Heckman [1979], with the only difference that the outcome variable is also binary, rather than being continuous. Adverse selection implies that the correlation between ν and η is positive. If we estimate a linear probability model for default, assuming that ν, η are bivariate normal with correlation coefficient ρ, we can employ the two step procedure of Heckman [1979] by first estimating a probit on d_i, and then constructing the Mills ratio and inserting it in the second equation. A test for a positive correlation between the error terms is a t-test on the coefficient of the Mills ratio in the default equation. As controls in the default equation we use firm level characteristics (total assets, share of intangible assets over total assets, returns on assets, leverage, sales, trade debit, score) as well as sector, year and area dummies. In the selection equation we add the indicators of local financial development in 1936 at the regional level collected by Guiso et al. [2004], who show

[15] As explained in the data section, we define a firm as defaulter if any of its loans are classified as bad up to at most 3 years after borrowing.

that they are good instruments for financial development today while uncorrelated with current economic performance. They therefore satisfy the condition for a valid exclusion restriction: they affect the probability of obtaining a loan, which varies with the degree of local financial development, but are unlikely to be correlated with the probability of defaulting, conditional on having a loan.[16]

Results reported in Table 2.2, Panel A, are consistent with the hypothesis that lending is affected by adverse selection. The coefficient of the Mills ratio is positive and statistically significant both when considering first loans and all loans. The magnitude is larger for the second sample, suggesting that adverse selection issues are not confined to the early phase of the firm's borrowing cycle.

2.3.2 Loan Size and Default

We then consider the relationship between amount of loan used and default probability. Differently from the previous subsection, we are not in a selection framework as the same firms are observed in both equations. Still, the idea is the same, as we test for a positive correlation between the unobservables that determine the choice of "coverage" and the occurrence of an "accident", conditional on several individual characteristics. We consider two dependent variables for coverage: the absolute amount of credit used as well as the amount of credit used as a share of credit granted. In our lending context we check if firms that use a larger share of their loans are more likely to default on them. Adverse selection should imply that riskier firms use more credit. We set up the following bivariate probit:

$$\begin{aligned} \ell_i &= \mathbf{1}(X_i\beta + \varepsilon_i > 0) \\ f_i &= \mathbf{1}(X_i\gamma + \eta_i > 0) \end{aligned} \tag{2.2}$$

where ℓ_i is a dummy equal to one if the loan amount used is above the median, or if the loan amount used over granted is above the median, and f_i takes value of one if the borrower is a defaulter. The vector of controls X_i is composed by year, area, sector, and bank fixed effects, as well as other firm's balance sheet variables, including the score, and the interest rate. We specify the distribution of the residuals ε_i, η_i as jointly normal, with a correlation coefficient ρ. Positive and significant ρ suggests presence of adverse selection. The results of this test are summarized in Table 2.2 Panel B. The positive correlation is similar for the sample of first loans and for all loans and for both dependent variables.

[16] This instrument is valid for this simplified setup of the reduced form test, but not for the structural model that we present later, where we need to instrument prices that vary at the bank-market-year level.

Table 2.2: Positive Correlation Tests

Panel A: Demand and Default		
	First	All
Selection	.131**	.312***
	(.059)	(.023)

Panel B: Loan Size and Default		
	First	All
Used	0.181***	0.170***
	(0.003)	(0.003)
Used/Granted	0.196***	0.186***
	(0.003)	(0.003)

Note: Panel A reports the selection term of a Heckman selection model. The two columns report the coefficient on the Mills ratio in a model where the outcome equation (default or not) is linear. Panel B reports the correlation coefficient of the error terms of a bivariate probit model. Columns labelled "First" only consider the first loan ever, "All" all loans.

2.4 The Model

The framework we construct aims at quantifying the effects of asymmetric information on the demand for and supply of credit for Italian firms. In order to test for this, we assume that each firm $i = 1, ..., I$ is willing to invest in a project and is looking for credit to finance it. Firms decide which bank $j = 1, ..., J$ to borrow from based on the conditions offered that maximise the expected "profits"[17] of their choice. This determines demand for credit. Conditional on demand, firms decide the amount of credit to use and whether to default or not. The supply of credit results from banks' static Bertrand-Nash competition on interest rates.

The theoretical model we develop is based on the following assumptions:

(1) **Asymmetric Information**: Following Stiglitz and Weiss [1981], we assume that the asymmetry of information is on the riskiness of the firm, known by the firm but not by the bank, whereas the distribution of riskiness among all firms is known by both. We identify this riskiness with the firm's probability of default. We let borrowers and lenders be risk neutral.[18]

(2) **First Year of New Loans**: We limit our analysis to the first year of newly granted loans. This is a common assumption in empirical models of insurance with asymmetric information, starting from Chiappori and Salanié [2000]. This is done to avoid heterogenous experience ratings among borrowers and loan renegotiation, as the focus of the paper is on first access to credit.[19]

(3) **Main New Credit Line**: We just consider the choice of the main new credit line that firms open for the first time within our sample. As shown by Detragiache et al. [2000], in Italy, multiple relationship banking is widely used by firms to reduce liquidity risk. However, the share of the main credit line opened accounts on average for over 70% of the total share of new yearly credit (both usable and used), justifying the choice of this simplifying assumption.

(4) **Posted Interest Rates**: We assume that banks have posted interest rates for types of firms $k = 1, ..., K$ in each market m and period t, depending on the borrowers' characteristics. Following the work by Albareto et al. [2011] on the determinants of interest rates decisions, these types are defined by the amount

[17] We will define these profits as utilities later on, to distinguish them from banks' profits.

[18] The assumption of asymmetric information in Stiglitz and Weiss [1981] is that lenders observe the mean return of a project, but not its riskiness.

[19] We relax this assumption in a companion paper (Pavanini and Schivardi [2013]).

of credit granted, the firm's sector, the firm's size in terms of sales, and the observable riskiness of the firm defined by the SCORE.[20]

(5) Exogenous Amount of Credit: We limit our analysis to the interest rate as the only screening device, as in Stiglitz and Weiss [1981]. Therefore, we assume that the amount of credit granted from bank j to firm i is exogenously given by the firm's project requirements, and that the bank just offers a posted interest rate for that specific amount to each type k in each market m. In a standard insurance or credit market with asymmetric information, firms are likely to compete not only on prices, but on other clauses of the contract as well. In our context, the amount granted could be another dimension over which banks compete. In a world with lending exclusivity, banks can offer menus of amounts granted with matched interest rates to reduce the extent of asymmetric information, for example charging rates that increase more than proportionally with the amount granted. However, this is the case only with contract exclusivity, which is not a feature of our setting, where borrowers can open multiple credit lines with different lenders. Empirical evidence of non-exclusivity results also from the pricing regression described in Appendix A, which presents a negative correlation between interest rates and the amount of credit granted.[21] Moreover, the assumption of setting the loan amount as part of the definition of type is also justified by the distribution of amounts granted, characterized by a high concentration of loans around some specific mass points. We also assume no collateral, as the type of loans we analyze are uncollateralized. We do however allow for an endogenous amount of loan used.

2.4.1 Demand, Loan Size and Default

Given these assumptions, let there be $i = 1, ..., I$ firms of observable type $k = 1, ..., K$ and $j = 1, ..., J$ banks in $m = 1, ..., M$ markets in period $t = 1, ..., T$. Let firms have the following utility from credit, which determines their demand:

$$U^D_{ikjmt} = \underbrace{\bar{\alpha}^D_0 + \alpha^D_1 P_{jmt} + X'_{jmt}\beta^D + \xi^D_{jmt}}_{\delta^D_{jmt}} + \underbrace{\sigma^D \nu_i + Y'_i \eta^D + \gamma^D_k}_{V^D_i} + \varepsilon^D_{ikjmt}. \tag{2.3}$$

[20] The construction of these posted interest rates is described in Appendix A.
[21] We thank Pierre-André Chiappori for his suggestions on this point.

We normalize to zero the utility from the outside option, which is not borrowing. Firms will choose the bank that maximizes their utility, or will choose not to borrow. Then, conditional on borrowing, they will choose the share of amount granted to use that maximizes the following utility:

$$U^L_{ikmt} = \underbrace{\alpha^L_0 + \alpha^L_1 P_{jmt} + X'_{jmt}\beta^L + \xi^L_{jmt}}_{\delta^L_{jmt}} + \underbrace{Y'_i\eta^L + \gamma^L_k}_{V^L_i} + \varepsilon^L_{ikmt}. \tag{2.4}$$

Finally, conditional on borrowing, they will choose to default if the following utility is greater than zero:

$$U^F_{ikmt} = \underbrace{\alpha^F_0 + \alpha^F_1 P_{jmt} + X'_{jmt}\beta^F + \xi^F_{jmt}}_{\delta^F_{jmt}} + \underbrace{Y'_i\eta^F + \gamma^F_k}_{V^F_i} + \varepsilon^F_{ikmt}. \tag{2.5}$$

Here X_{jmt} are banks' observable attributes, P_{jmt} are the posted interest rates mentioned above,[22] ξ_{jmt} are banks' unobservable (to the econometrician) attributes, Y_i are firms' observable characteristics, and γ_k are types' fixed effects. We assume that ε^D_{ikjmt} is distributed as a type 1 extreme value, following the literature on demand estimation for differentiated products (Berry [1994], Berry et al. [1995]). We let the random coefficient of the demand's constant term $\alpha^D_{0i} = \bar{\alpha}^D_0 + \sigma^D\nu_i$, with $\nu_i \sim N(0,1)$,[23] to be jointly normally distributed with ε^L_{ikmt}, and ε^F_{ikmt}, such that:

$$\begin{pmatrix} \alpha^D_0 \\ \varepsilon^L \\ \varepsilon^F \end{pmatrix} \sim N\left(\begin{pmatrix} \bar{\alpha}^D_0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^{2D} & \rho_{DL}\sigma^D\sigma^L & \rho_{DF}\sigma^D\sigma^F \\ \rho_{DL}\sigma^D\sigma^L & \sigma^{2L} & \rho_{LF}\sigma^L\sigma^F \\ \rho_{DF}\sigma^D\sigma^F & \rho_{LF}\sigma^L\sigma^F & \sigma^{2F} \end{pmatrix} \right). \tag{2.6}$$

We interpret a positive correlation between the firm specific unobservables driving demand and default (ρ_{DF}) as evidence of adverse selection. The intuition is that if the unobservables that drive demand are positively correlated with the unobservables that drive default, then riskier firms are more likely to demand. The idea behind the identification of the correlation between α^D_0 and ε^F is the following. If we observe a firm taking out a loan, while the model tells us that this firm should be unlikely to take the loan, then this is a "high α^D_0" firm. A positive correlation of

[22] We explain in Appendix A how we separate the bank-market-period price from the type specific price.

[23] We use 100 Halton draws for simulation. According to Train and Winston [2007], 100 Halton draws achieve greater accuracy in mixed logit estimations than 1,000 pseudo-random draws.

α_0^D with ε^F is evidence of adverse selection.

We interpret a positive correlation between the unobservables driving loan size and default (ρ_{LF}) as other possible evidence of adverse selection. The intuition is that if the unobservables that drive the choice of how much credit to use are positively correlated with the unobservables that drive default, then riskier firms will use more credit. With this definition of adverse selection we are trying to capture the case in which a risky firm (high ε^F), before signing the contract, already knows that due to its high ε^L it will use a higher share of the loan. However, our definition cannot rule out the case in which two ex-ante equally risky firms take the same loan, and one of them is hit by a negative shock after the contract has been signed. This shock increases ε^L for the firm that was hit, forcing it to use more of the loan.[24] This identification strategy allows us to recover adverse selection parameters that are common across banks and markets, not bank or market specific.[25]

This set up is similar to Einav et al. [2012], but differs in the specification of the demand utility. In our case, borrowers' choices follow a multinomial distribution, instead of a binomial. This raises the issue of correlating residuals from the demand model, which vary across borrowers and alternatives (i.e. lenders), to the residuals from the loan size and default models, which instead vary only across borrowers. We follow the approach of Ackerberg and Botticini [2002] and allow the normally distributed random coefficient on the constant term to be correlated with the residuals from the loan size and default equations. We argue that this a practical and intuitive solution, as it simplifies the problem and allows for a correlation between unobservables only at the level of the borrower. This implies that in the presence of adverse selection a riskier firm is more likely to demand from any lender, and not differently across different lenders.

2.4.2 Supply

On the supply side, we let banks set their interest rates competing à la Bertrand Nash. We assume that bank j's profits in market m at time t are given by the sum of the profits made with each subset of its borrowers of types k:

$$\begin{aligned} \Pi_{jkmt} &= (P_{jkmt} - MC_{jmt})Q_{jkmt}(1 - F_{jkmt}) - MC_{jmt}Q_{jkmt}F_{jkmt} \\ &= P_{jkmt}Q_{jkmt}(1 - F_{jkmt}) - MC_{jmt}Q_{jkmt}, \end{aligned} \tag{2.7}$$

[24] In this case, ρ_{LF} could be interpreted as evidence of either adverse selection or moral hazard. See Abbring et al. [2003] for distinguishing between those sources of asymmetric information.

[25] There is not a clear economic interpretation of the correlation between the demand and loan size unobservables, so at the moment we set it to zero for simplicity.

where Q_{jkmt} and F_{jkmt} are bank's expectation of demand and default. In particular, Q_{jkmt} is given by the model's market shares and the expected loan size, and F_{jkmt} is the average default rate for the borrowers of type k that bank j lends to in market m. P_{jkmt} is the price of the loan $(1 + r_j)$. MC_{jmt} are the bank's marginal costs, which we assume to be constant at the bank-market-period level. It is important to note that F_{jkmt} depends on price through two channels. First, equation (2.5) allows for a direct impact of the interest rate on default probability. Second, a higher interest rate also changes the composition of borrowers as stated in Assumption 1: increasing price increases the conditional expectation of α_0^D, as safer firms are more likely to self-select out of the borrowing pool. If $\rho_{DF} \geq 0$, this implies that an increase in prices increases the probability of default of the pool of borrowers.

The first order conditions of this profit function deliver the following pricing equation:

$$P_{jkmt} = \underbrace{\frac{MC_{jmt}}{1 - F_{jkmt} - F'_{jkmt}\frac{Q_{jkmt}}{Q'_{jkmt}}}}_{\textit{Effective Marginal Costs}} - \underbrace{\frac{(1 - F_{jkmt})\frac{Q_{jkmt}}{Q'_{jkmt}}}{1 - F_{jkmt} - F'_{jkmt}\frac{Q_{jkmt}}{Q'_{jkmt}}}}_{\textit{Markup}}, \tag{2.8}$$

Note that the equilibrium price depends on what we define as "effective" marginal costs and on a markup term. F'_{jkmt} is the derivative of the expected default rate with respect to prices, and Q'_{jkmt} is the derivative of the market share with respect to prices. $\frac{Q_{jkmt}}{Q'_{jkmt}}$ would be the markup in a Bertrand-Nash model with differentiated products and no asymmetric information. In fact if there was no default, i.e. $F_{jkmt} = F'_{jkmt} = 0$, we would be back to a standard equilibrium pricing equation for differentiated firms competing à la Bertrand-Nash as in Berry et al. [1995]. We will analyze this equilibrium pricing equation in greater detail in the next section.

2.4.3 Monte Carlo

We construct a simple numerical example to give the intuition underlying the model's predictions. We simulate data for the case of a monopoly bank facing $i = 1, ..., N$ heterogeneous borrowers. For simplicity, we concentrate on adverse selection between demand and default (ρ_{DF}), setting loan size to 1 and $\rho_{DL} = \rho_{LF} = 0$. We keep this data fixed and vary the number of banks, borrowers' price sensitivity, and the extent of asymmetric information, where $\rho_{DF} < 0$ means advantageous selection

and $\rho_{DF} > 0$ means adverse selection. For each of these cases we compute banks' equilibrium prices based on our model. Let borrower i have U_{ij}^D utility from taking credit from bank j, U_{i0}^D utility from not borrowing, and U_i^F utility from defaulting:

$$\begin{aligned} U_{ij}^D &= \alpha_{0i} + \alpha_1 P_j + \epsilon_{ij}, \\ &= \bar{\alpha}_0 + \sigma\nu_i + \alpha_1 P_j + \epsilon_{ij}, \\ U_{i0}^D &= \epsilon_{i0}, \\ U_i^F &= \varepsilon_i, \end{aligned} \tag{2.9}$$

where P_j is the interest rate charged by bank j, ϵ_{ij}, ϵ_{i0} are distributed as type 1 extreme value, and $\nu_i \sim N(0,1)$. We set $\sigma = 2$ and $\bar{\alpha}_0 = 1$, and allow α_i and ε_i to be jointly normally distributed, with correlation coefficient $-1 \leq \rho \leq 1$. We assume that all the borrowers have the same price sensitivity $\alpha_1 < 0$. Our asymmetric information assumption implies that a bank doesn't observe its borrowers' individual default probability, but only its distribution. As a consequence, it only offers one pooling price P_j for everyone. Given this setup, the demand probability will be given by:

$$\begin{aligned} \Pr_{ij}^D &= \Pr(\alpha_{0i} + \alpha_1 P_j + \epsilon_{ij} > \alpha_{0i} + \alpha_1 P_h + \epsilon_{ih} \ \forall h \neq j) \\ &= \tfrac{\exp(\alpha_{0i}+\alpha_1 P_j)}{1+\sum_\ell \exp(\alpha_{0i}+\alpha_1 P_\ell)} \\ &= \Lambda(\alpha_{0i} + \alpha_1 P_j), \end{aligned} \tag{2.10}$$

and we will construct banks' market shares as $S_j = \frac{1}{N}\sum_i \Pr_{ij}^D$. Conditional on demand, default probability will follow from Wooldridge [2002] as:

$$\begin{aligned} \Pr_{ij}^{F|D=j} &= E\big[\Pr(F=1|\nu, P_j)|D=j, P_j\big] \\ &= \tfrac{1}{\Lambda(\alpha_{0i}+\alpha_1 P_j)} \int_{-(\alpha_{0i}+\alpha_1 P_j)}^{\infty} \Phi\left(\frac{\frac{\rho\nu}{\sigma^2}}{\sqrt{1-\frac{\rho^2}{\sigma^2}}}\right)\phi(\nu)d\nu, \end{aligned} \tag{2.11}$$

and we will construct banks' share of defaulters as $F_j = \frac{1}{N_j}\sum_i \Pr_{ij}^F$, where N_j is the number of borrowers of bank j. Given these probabilities and our supply side model described in equations (3.9) and (3.10), the first order conditions will deliver the following equilibrium pricing equation for each bank:

$$P_j^* = \underbrace{\frac{MC}{1 - F_j - F_j' \frac{1}{\alpha_1(1-S_j)}}}_{\textit{Effective Marginal Costs}} - \underbrace{\frac{(1-F_j)\frac{1}{\alpha_1(1-S_j)}}{1 - F_j - F_j' \frac{1}{\alpha_1(1-S_j)}}}_{\textit{Markup}}, \tag{2.12}$$

where the first term on the right hand side represents what we define as "effective" marginal costs (EffMC), and the second term represents the markup

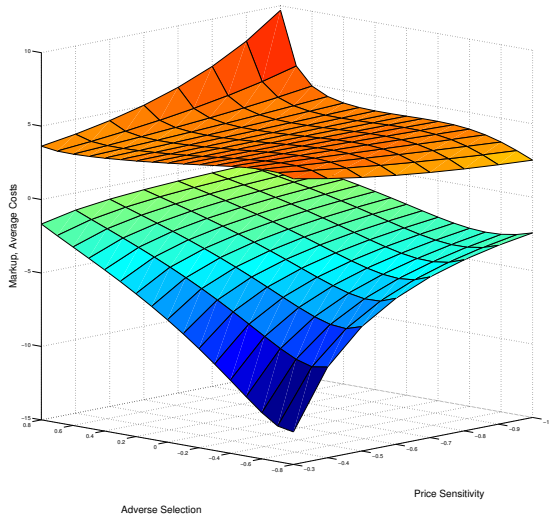
(MKP). F_j' is the derivative of the expected default rate with respect to prices, and $\alpha_1(1 - S_j)$ is the derivative of the market share with respect to prices. For $\alpha_1 < 0$ it can be shown that the EffMC term is always positive and the markup term is always negative.

The different effects of these two factors on equilibrium prices is crucial to understand the interaction between asymmetric information and imperfect competition. This is displayed in Figures 2.1 and 2.2, where the top graph represents EffMC above and negative of the markup below, and the bottom graph shows equilibrium prices for a monopolist bank. We let these three elements vary across different degrees of adverse selection, measured by ρ, and competition, measured by α_1. This means that for the moment we are capturing competition versus the outside option, but we have verified that increasing the number of banks gives the same result.

Looking at Figure 2.1, for a high level of competition (i.e. rightmost point on the figure) an increase in adverse selection (moving to the northwest) causes EffMC to increase, whereas for low competition (point closest to the reader, again moving northwest) they remain relatively constant. The intuition for this result is the following. Higher adverse selection implies higher correlation between borrowers' willingness to pay (WTP) and their riskiness. Hence, with strong competition only firms with high WTP will borrow, whereas with less competition even firms with low WTP will take credit, lowering the average riskiness of the pool of borrowers. The opposite happens for the markup curve as we increase adverse selection, because it remains nearly constant for high competition (leftmost point, moving to the northeast), but it decreases substantially for a low level of competition (closest point to the reader, moving to the northeast). What the graph shows in fact is that both an increase in adverse selection and an increase in competition reduces a bank's markup, implying that adverse selection has a mitigating effect on market power.

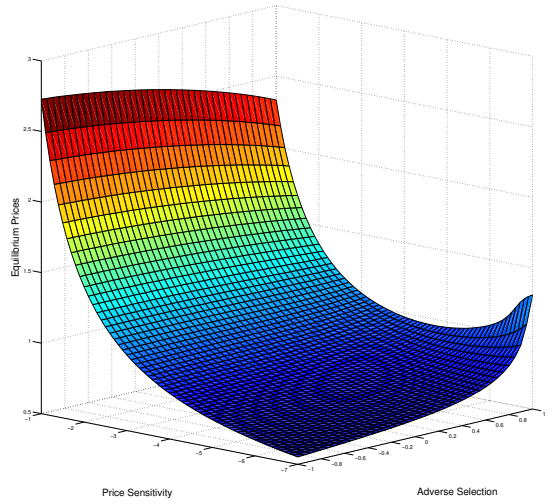
As shown in Figure 2.2, the combination of these two factors results in a non-monotonic equilibrium price response to an increase in adverse selection. If on one hand equilibrium prices rise in a very competitive environment (closest point to the reader, moving to the northeast), the opposite happens in a concentrated market (leftmost point, moving to the northeast). This is because in the first case the increasing EffMC drive prices up, whereas in the second case the declining markup drives prices down. More intuitively, in a highly competitive market where banks have small price-cost margins, higher prices is the only possible response to an increase in adverse selection. However, banks with a higher price-cost margin will find it more profitable to reduce prices, as this will allow them to lower the average riskiness to their pool of borrowers.

Figure 2.1: Adverse Selection vs Imperfect Competition - Effective Marginal Costs, negative Markups



Note: The vertical axis shows the value of effective marginal costs and of the negative of the markup. The left horizontal axis is level of adverse selection, increasing towards left. The right horizontal axis is the level of price sensitivity (our measure of competition with the outside option), increasing towards the right.

Figure 2.2: Adverse Selection vs Imperfect Competition - Equilibrium Prices



Note: The vertical axis shows the level of equilibrium prices. The left horizontal axis is level of price sensitivity (our measure of competition with the outside option), increasing towards the right. The right horizontal axis is the level of adverse selection, increasing towards right. The axis definitions in this figure differ from those in Figure 2.1 to better display the effects in each.

2.5 Econometric Specification

Following the model presented above, let $m = 1, ..., M$ index a province, $t = 1, ..., T$ a year, $i = 1, ..., I$ the firm that borrows, and $j = 1, ..., J_{mt}$ be the bank/loan identifier in market m at time t. Moreover, let $k = 1, ..., K$ identify the type of firm that is borrowing. The k index further segments the market, as banks can lend across all types of firms within the same market, but firms can only borrow at the interest rate offered to their own type. Let Y_i be a vector of firm and firm-bank specific characteristics (firm's balance sheet data, firm's age, and firm's distance to the closest branch of each bank), X_{jmt} a vector of bank-province-year specific attributes (number of branches in the market, years of presence in the market, bank fixed effects), and γ_k types' fixed effects.

We estimate a system of three equations: demand for credit, amount of loan used, and default. We use a 2-step method based on maximum simulated likelihood and instrumental variables (Train [2009]). In the first step we estimate the firm-level parameters $\eta = \{\eta^D, \eta^L, \eta^F\}$, the types' fixed effects $\gamma_k = \{\gamma_k^D, \gamma_k^L, \gamma_k^F\}$, the correlation coefficients $\rho = \{\rho_{DF}, \rho_{DL}, \rho_{LF}\}$, and the covariances σ^D and σ^L from the firms' choice probabilities.[26] We follow Einav et al. [2012], but differ from them as we estimate demand using a mixed logit with random coefficients, rather than a probit. We also recover the lender-province-year specific constants $\widehat{\delta}_{jmt} = \{\widehat{\delta}_{jmt}^D, \widehat{\delta}_{jmt}^L, \widehat{\delta}_{jmt}^F\}$ using the contraction method introduced by Berry [1994].

The probability that borrower i of type k in market m at time t chooses lender j is given by:

$$\Pr_{ikjmt}^D = \int \left[\frac{\exp(\widehat{\delta}_{jmt}^D(X_{jmt}, P_{jmt}, \xi_{jmt}^D, \beta^D) + V_i^D(Y_i, \eta^D, \gamma_k^D))}{1 + \sum_\ell \exp(\widehat{\delta}_{\ell mt}^D(X_{\ell mt}, P_{\ell mt}, \xi_{\ell mt}^D, \beta^D) + V_i^D(Y_i, \eta^D, \gamma_k^D))} \right] f(\alpha_{0i}^D|\theta) d\alpha_{0i}^D, \tag{2.13}$$

where $f(\alpha_{0i}^D|\theta)$ is the density of α_{0i}^D, and θ are the parameters of its distribution that we want to estimate. The estimation of this choice model only provides the estimates of $\eta^D, \gamma_k^D, \sigma^D$, but not of the parameters in δ^D. Looking at the second equation, the share of credit used over granted conditional on borrowing, the

[26] In this version of the paper we are still not estimating ρ_{DL}, which we set to zero, and σ^D, which we set to 1. For the second, it is due to the well known identification problem of the standard deviations of random coefficients in Berry et al. [1995], explained in Berry et al. [2004] and Train and Winston [2007]. We are working on incorporating second preferred choices into the model to guarantee better identification and to be able to estimate this parameter.

probability of observing a utilization of L_{ikmt} is given by:

$$\begin{aligned}\Pr^{L}_{ikmt,L=L^*|D=1,\alpha^D_{0i}} &= \Pr(L_{ikmt} = \delta^L_{jmt} + V^L_i + \varepsilon^L_{ikmt}|\alpha^D_{0i}) \\ &= \int \frac{1}{\tilde{\sigma}_{\varepsilon^L_{ikmt}|\alpha^D_{0i}}} \phi_{\varepsilon^L_{ikmt}|\alpha^D_{0i}} \left(\frac{L_{ikmt} - \hat{\delta}^L_{jmt}(X_{jmt}, P_{jmt}, \xi^L_{jmt}, \beta^L) - V^L_i(Y_i, \eta^L, \gamma^L_k) - \tilde{\mu}_{\varepsilon^L_{ikmt}|\alpha^D_{0i}}}{\tilde{\sigma}_{\varepsilon^L_{ikmt}|\alpha^D_{0i}}} \right) f(\alpha^D_{0i}|\theta) d\alpha^D_{0i},\end{aligned} \tag{2.14}$$

$$\text{where} \qquad \varepsilon^L_{ikmt}|\alpha^D_{0i} \sim N\big(\underbrace{\sigma^L \rho_{DL} \nu_i}_{\tilde{\mu}_{\varepsilon^L_{ikmt}|\alpha^D_{0i}}}, \underbrace{\sigma^{2L}(1-\rho^2_{DL})}_{\tilde{\sigma}^2_{\varepsilon^L_{ikmt}|\alpha^D_{0i}}}\big)$$

where ϕ is a standard normal pdf. Finally, the probability of default conditional on taking a loan is:

$$\Pr^{F}_{ikmt,F=1|D=1,\alpha^D_{0i},\varepsilon^L_{ikmt}} = \int \Phi_{\varepsilon^F_{ikmt}|\alpha^D_{0i},\varepsilon^L_{ikmt}} \left[\frac{\hat{\delta}^F_{jmt}(X_{jmt}, P_{jmt}, \xi^F_{jmt}, \beta^F) + V^F_i(Y_i, \eta^F, \gamma^F_k) - \tilde{\mu}_{\varepsilon^F_{ikmt}|\alpha^D_{0i},\varepsilon^L_{ikmt}}}{\tilde{\sigma}_{\varepsilon^F_{ikmt}|\alpha^D_{0i},\varepsilon^L_{ikmt}}} \right] f(\alpha^D_{0i}|\theta) d\alpha^D_{0i}, \tag{2.15}$$

$$\text{where} \qquad \varepsilon^F_{ikmt}|\alpha^D_{0i}, \varepsilon^L_{ikmt} \sim N\big(\underbrace{A\sigma^D \nu_i + B\varepsilon^L_{ikmt}}_{\tilde{\mu}_{\varepsilon^F_{ikmt}|\alpha^D_{0i},\varepsilon^L_{ikmt}}}, \underbrace{\sigma^{2F} - (A\rho_{DF} + B\rho_{LF})}_{\tilde{\sigma}^2_{\varepsilon^F_{ikmt}|\alpha^D_{0i},\varepsilon^L_{ikmt}}}\big)$$

$$A = \frac{\rho_{DF}\sigma^{2L} - \rho_{LF}\rho_{DL}}{\sigma^{2D}\sigma^{2L} - \rho^2_{DL}}$$

$$B = \frac{-\rho_{DF}\rho_{DL} + \rho_{LF}\sigma^{2D}}{\sigma^{2D}\sigma^{2L} - \rho^2_{DL}}$$

where the residuals ε^F_{ikmt} are conditional on demand and loan amount unobservables. Similarly to the demand side, the estimation of these two choice equations, jointly with the demand one, only delivers the parameters $\eta^L, \eta^F, \gamma^L_k, \gamma^F_k, \rho, \sigma^L$.

In the second step, the estimated constants $\widehat{\delta}_{jmt}$ are the dependent variables of instrumental variable regressions that recover the parameters $\bar{\alpha}^D_0, \alpha^D_1, \alpha^L_0, \alpha^L_1, \alpha^F_0, \alpha^F_1, \beta^D, \beta^L, \beta^F$ of the bank specific attributes X_{jmt} and prices P_{jmt}. This second step also controls for the potential endogeneity bias caused by the correlation between prices and unobserved (to the econometrician) bank attributes $\xi_{jmt} = \{\xi^D_{jmt}, \xi^L_{jmt}, \xi^F_{jmt}\}$. Following Berry [1994], the contraction method on the demand side finds the δ^D that equate predicted market shares $\widehat{S}^D_{jmt}$ to actual market shares S^D_{jmt}. This iterative process is

defined by:

$$\delta_{jmt}^{D,r+1} = \delta_{jmt}^{D,r} + \ln\left(\frac{S_{jmt}^{D}}{\widehat{S}_{jmt}^{D}(\delta_{jmt}^{D,r})}\right). \tag{2.16}$$

The predicted market shares are defined as $\widehat{S}_{jmt}^{D} = \sum_i \Pr_{ikjmt}^{D}/N_{mt}$, where N_{mt} are the number of borrowers in market m at time t. Given the value of these constant terms, the parameters $\bar{\alpha}_0^D, \alpha_1^D, \beta^D$ are estimated using instrumental variables:

$$\delta_{jmt}^{D} = \bar{\alpha}_0^D + \alpha_1^D P_{jmt} + X'_{jmt}\beta^D + \xi_{jmt}^{D}, \tag{2.17}$$

with ξ_{jmt}^{D} being the mean zero structural econometric error term. Similarly, the lender-market constants for loan size δ_{jmt}^{L} and default δ_{jmt}^{F} are estimated using a nonlinear least squares search routine as in Goolsbee and Petrin [2004], which solves for:

$$\delta_{jmt}^{L} = \arg\min_{\delta} \sum_j \left(\hat{S}_{jmt}^{L}(\eta^L, \delta^L) - S_{jmt}^{L}\right)^2, \tag{2.18}$$

$$\delta_{jmt}^{F} = \arg\min_{\delta} \sum_j \left(\hat{S}_{jmt}^{L}(\eta^F, \delta^F) - S_{jmt}^{L}\right)^2, \tag{2.19}$$

where $\widehat{S}_{jmt}^{L}, \widehat{S}_{jmt}^{F}$ and S_{jmt}^{L}, S_{jmt}^{F} are the predicted and actual shares of loan sizes and defaults for lender j in market m at time t. Given the value of these constant terms, the parameters $\alpha_0^L, \alpha_1^L, \beta^L$ and $\alpha_0^F, \alpha_1^F, \beta^F$ are estimated using instrumental variables:

$$\delta_{jmt}^{L} = \alpha_0^L + \alpha_1^L P_{jmt} + X'_{jmt}\beta^L + \xi_{jmt}^{L}, \tag{2.20}$$

$$\delta_{jmt}^{F} = \alpha_0^F + \alpha_1^F P_{jmt} + X'_{jmt}\beta^F + \xi_{jmt}^{F}. \tag{2.21}$$

2.6 Estimation

Following from section 5, we use the demand, loan size and default probabilities to construct the simulated maximum likelihood that allows us to recover the parameters in $\eta, \gamma_k, \sigma^D, \sigma^L, \rho$:

$$\log L = \sum_i \log(\Pr^D_{ikjmt}) d_{ikjmt} + \sum_{i \in D} \Big[\log(\Pr^L_{ikmt}) + \log(\Pr^F_{ikmt}) f_{ikmt} + \log(1 - \Pr^F_{ikmt})(1 - f_{ikmt}) \Big], \tag{2.22}$$

where d_{ikjmt} is the dummy for the choice by firm i of type k of bank j in market m at time t, and f_{ikmt} is the dummy identifying its default. In order to estimate the remaining parameters we need an additional step explained below.

2.6.1 Constructing the Sample

As already mentioned, we focus on the first line of credit that a firm opens (at least within our dataset), excluding the first year (1988). We do this to concentrate on new borrowers, where we expect to find stronger asymmetric information, and because modeling the evolution of the borrower-lender relationship is beyond the scope of this paper.[27] [28] Following other papers on Italian local credit markets (Felici and Pagnini [2008], Bofondi and Gobbi [2006], Gobbi and Lotti [2004]), we identify banking markets as the Italian provinces, also used by Italian supervisory authorities as proxies for the local markets for deposits.[29] Our markets are then constructed as province-year combinations. We define the loan size variable as the share of loan used over loan granted, and define default as an ever default variable, as explained before.

The observable explanatory variables that determine firm's demand, loan size and default choices are firm and bank characteristics, summarized in Table 2.1. In the first set of regressors we include firms' fixed assets, the ratio of intangible over total assets, net worth, trade debit, profits, cash flow, and age, where trade debit is the debit that the firm has with its suppliers or clients. We also include types' fixed effects, where a type is defined as a combination of amount granted, sector, size, and score.[30] In the second group we use prices, bank's share of branches in the

[27] We do this in a companion paper Pavanini and Schivardi [2013].

[28] A more extensive description of the construction of the sample is in Appendix A.

[29] See Pavanini and Schivardi [2013] and Ciari and Pavanini [2013] for a detailed discussion on the definition of local banking markets in Italy.

[30] See the Appendix A for a detailed description of the types.

province, number of years the bank had at least one branch in the province, and bank dummies. We also control for the distance between each firm and the closest branch of each bank. We provide details on these variables in the appendix. We motivate the choice of these explanatory variables in Section 6.3.

2.6.2 Identification

The use of instrumental variables in the second step of the estimation aims at correcting the potential endogeneity bias in the price coefficient for the three equations. The bias derives from the possible correlation between prices P_{jmt} and unobserved (to the econometrician) bank-market level characteristics ξ_{jmt}. These unobserved attributes can be thought as the borrowers' valuation of a banks' brand, quality, and credibility, which are assumed to influence borrowers' demand, loan size, and default decisions, but are also very likely to be correlated with banks' interest rates. Think for example of ξ_{jmt} as a banks' reputation for offering valuable and helpful assistance to its borrowers in their business projects, which is unobserved to the econometrician. Borrowers will value this quality when deciding which bank to get credit from, and they will also be affected in their likelihood of using more or less credit and of defaulting. Consequently, the bank will be likely to charge a higher interest rate, given the potentially higher markup that this attribute can provide. Moreover, assuming default is increasing in interest rates, a good assistance can lower the borrower's default probability, allowing banks to charge a higher rate.

To address the simultaneity problem, following Nevo [2001], we include bank dummies to capture the bank characteristics that do not vary by market (year-province). This means that the correlation between prices and banks' nationwide-level unobserved characteristics is fully accounted for with these fixed effects, and does not require any instruments. Hence, we can rewrite equation (2.17), and similarly equations (2.20) and (2.21), as:

$$\delta^D_{jmt} = \bar{\alpha}^D_0 + \alpha^D_1 P_{jmt} + X'_{jmt}\beta^D + \xi^D_j + \Delta\xi^D_{jmt}, \tag{2.23}$$

where ξ^D_j are banks' fixed effects and $\Delta\xi^D_{jmt}$ are bank-market-time specific deviations from the national mean valuation of the bank. Therefore, we need to use instrumental variables to account for the potential correlation between interest rates and these bank-market-time specific deviations. We argue that a valid instrument is represented by the share of branches in a specific market of merging rival banks.[31] Since mergers only happen in a single year, this accounts to relying on the

[31] We experimented also with other instruments, with similar results. In one case we followed the

across time correlation of prices with changes in concentration among branches at the market level. The first stage regression shows that this correlation is positive and significant, implying that greater rivals' concentration leads to higher interest rates. We verify empirically the rank condition for instruments' validity with the first stage estimates[32], showing that the instruments are good predictors of interest rates. We compare OLS and IV second stages, to show how the instruments lessen the simultaneity bias.[33] Last, for the exclusion restriction to hold, we assume that bank-market-time specific deviations $\Delta\xi^D_{jmt}$ are uncorrelated with the share of branches of merging rival banks in a market-time combination. We interpret these deviations, for example, as market specific differences in a bank's quality with respect to its national average quality. These can be thought as differences in local managers' capacities, or in a bank's management connections with the local industries and authorities. These factors are likely to influence a bank's prices in that local market, but not the merging decision of rivals, which are usually taken at a national level and are effective across various markets.

2.6.3 Results

The estimates of the structural model are presented in Table 2.3. The three columns of results refer respectively to the demand, loan size and default equations. The top part of the table shows the effect of firm characteristics, the middle one the effect of bank characteristics, and the bottom one shows the correlation coefficients of interest, i.e. the correlation between unobservables of demand and default (ρ_{DF}) and the correlation between unobservables of loan size and default (ρ_{LF}). We decided to include those specific firm characteristics to control for different measures of firms' assets, profitability, debt, age, and distance, and for our definition of observable type. We chose among the wide set of balance sheet variables running various reduced form regressions for demand, loan size, and default. We wanted to control for different measures of firm size, in the form of assets[34], but also for some measures of firms' current performance, in terms of profits and cash flow. We also tried to control for

approach of Nevo [2001] and Hausman and Taylor [1981], which implies instrumenting the prices charged by a bank j in a market m with the average of the prices that the same bank charges in all the other markets. We also tried with banks' expenditure in software per employee, weighted by the number of branches in a market, and with the sum of rivals' characteristics, as in Berry et al. [1995].

[32] First stage estimates are reported in Appendix B.

[33] OLS and IV second stage estimates are reported in Appendix B.

[34] Albareto et al. [2011] describe the importance of firms' size in the organization of lending in the Italian banking sector.

other specific forms of finance that firms have access to, such as debt from suppliers[35]. Finally, we computed the firm's age and the distance between the city council where the firm is located and the city council where the closest branch of each bank in the firm's choice set is located.[36] As introduced in the previous sections, we include fixed effects for the type of the firm as a determinant of the posted prices.[37] Following the survey of Albareto et al. [2011], these types are constructed as the combination of the firm's sector (primary, secondary, tertiary), size (sales above or below the median), riskiness (three risk categories based on the SCORE), and amount granted (five categories between 0 and 3,000,000 €). We also included the number and the share of branches that a bank has in a market (province-year), as well as the number of years that it has been in the market. We have data on branches from 1959, so we can observe banks' presence in each council for the 30 years before the beginning of our loan sample. These variables aim at capturing the level of experience that a bank has in a market, as well as the density of its network of branches with respect to its competitors, which can both be relevant features influencing firms' decisions.

The estimates present evidence of asymmetric information, both in terms of the correlation between demand and default unobservables and loan size and default unobservables. This confirms the results of the the reduced form test that we presented earlier. Looking at the demand side, we find that distance and prices have a negative impact on demand, as expected. In general, it seems that firms with more net worth, cash flow and trade debit are less likely to demand credit, but firms with more fixed assets and profits are more likely to borrow. Older firms are also less likely to demand. Firms seem to favor banks with a higher share of branches, but are less likely to demand form banks with longer experience in a market. This might be because the sample we are considering is of new borrowers, which might be perceived as more risky by experienced banks. Hence, these firms are more likely to get better conditions from less experienced banks. The share of loan used seems to follow the same logic as demand for fixed assets and ratio of intangible assets, as well as profits, cash flow, interest rates and share of branches. Differently from demand, the share of loan used over granted is increasing in the distance from the branch firms are borrowing from. For what concerns the default probability, this

[35] Petersen and Rajan [1995] use the amount of trade credit as a key variable to determine if borrowers are credit constrained, as it's typically a more expensive form of credit than banks' credit lines.

[36] It is important to include distance as Degryse and Ongena [2005] show empirical evidence, using Belgian data, that in lending relationship transportation costs cause spatial price discrimination. They find that loan rates decrease with the distance between the borrower and the lender, and increase with the distance between the borrower and the competing lenders.

[37] See Appendix A for a detailed description of how we construct types.

is negatively influenced by more assets, cash flow and trade debit, but positively affected by net worth, profits, and firm's age. As expected, higher interest rates increase default probability.

The mean of own and cross price elasticities for the main 5 banks in the sample are reported in Table 2.4. We find that on average a 1% increase in interest rate reduces a bank's own market share by over 2%, and increases competitor banks' shares by about 0.2%.

Table 2.3: Structural Estimates

Variables			Demand	Loan Size	Default
1st Stage Firm Level	**Assets**	Fixed Assets	0.197***	0.018**	-0.026**
			(0.017)	(0.007)	(0.009)
		Intangible/Total Assets	1.352***	2.001***	-1.189***
			(0.429)	(0.123)	(0.211)
		Net Worth	-0.370***	0.021	0.116***
			(0.024)	(0.016)	(0.016)
	Profitability	Profits	0.623***	0.040***	0.167***
			(0.036)	(0.016)	(0.014)
		Cash Flow	-0.327***	-0.049***	-0.451***
			(0.028)	(0.015)	(0.027)
	Debt	Trade Debit	-0.692***	0.003	-0.116***
			(0.044)	(0.018)	(0.023)
	Others	Firm's Age	-0.798***	-0.036	0.131***
			(0.050)	(0.029)	(0.014)
		Distance	-5.481***	0.088**	-0.020
			(0.257)	(0.037)	(0.019)
		Type FE	Yes	Yes	Yes
2nd Stage Bank Level		Interest Rate	-3.669***	-0.295***	2.387***
			(0.348)	(0.097)	(0.389)
		Number of Branches	-2.746***	-0.102	-0.236
			(0.269)	(0.075)	(0.300)
		Share of Branches	12.646***	0.429**	0.098
			(0.721)	(0.201)	(0.805)
		Years in Market	-1.001***	-0.032	-0.124
			(0.172)	(0.048)	(0.193)
		Bank FE	Yes	Yes	Yes
Adverse Selection	**Demand-Default**	ρ_{DF}			0.304***
					(0.006)
	Loan Size-Default	ρ_{LF}			0.159***
					(0.001)
Obs			301,334	7,170	7,170

Note: Standard errors in brackets. $*$ is significant at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

Table 2.4: Mean across Markets of Own and Cross Price Elasticities for Main Banks

Banks	**Bank 1**	**Bank 2**	**Bank 3**	**Bank 4**	**Bank 5**
Bank 1	-2.635	0.133	0.206	0.168	0.136
Bank 2	0.164	-2.394	0.163	0.164	0.154
Bank 3	0.261	0.258	-2.662	0.276	0.257
Bank 4	0.202	0.171	0.234	-2.562	0.162
Bank 5	0.184	0.161	0.183	0.174	-2.734

Note: These are the first 5 banks in terms of national market shares. Elasticities are interpreted as the percentage change in market shares in response to a 1% increase in prices.

2.7 Counterfactuals

We run a counterfactual policy experiment to quantify the effects of asymmetric information, as well as to understand the relationship between asymmetric information and imperfect competition. We simulate an increase in adverse selection, and analyze the consequence of this change on equilibrium prices, quantities, and defaults. An increase in adverse selection captures the idea that during a financial crisis investment opportunities contract for all firms, but risky firms will be more exposed than safer ones, demanding more credit. We simulate this scenario doubling the estimated correlation coefficients. Once we recover the new equilibrium outcomes of interest in the new scenario, we investigate whether the variations that we observe from the baseline model are correlated with various measures of competition in the different local markets.

In this counterfactual exercise we follow the example of Nevo [2000] and recover each bank's marginal costs using the pricing equation (3.10):

$$\widehat{MC}_{jmt} = P_{jmt}\Big[1 - F_{jmt} - F'_{jmt}\frac{Q_{jmt}}{Q'_{jmt}}\Big] + \frac{(1 - F_{jmt})\frac{Q_{jmt}}{Q'_{jmt}}}{1 - F_{jmt} - F'_{jmt}\frac{Q_{jmt}}{Q'_{jmt}}} \tag{2.24}$$

Under the assumption of marginal costs being the same in each scenario, we re-calculate banks' market shares, loan sizes and defaults with the counterfactual level of adverse selection, and derive the new equilibrium prices as:

$$\widetilde{P}_{jmt} = \frac{\widehat{MC}_{jmt}}{1 - \widetilde{F}_{jmt} - \widetilde{F}'_{jmt}\frac{\widetilde{Q}_{jmt}}{\widetilde{Q}'_{jmt}}} - \frac{(1 - \widetilde{F}_{jmt})\frac{\widetilde{Q}_{jmt}}{\widetilde{Q}'_{jmt}}}{1 - \widetilde{F}_{jmt} - \widetilde{F}'_{jmt}\frac{\widetilde{Q}_{jmt}}{\widetilde{Q}'_{jmt}}}, \tag{2.25}$$

where $\widetilde{Q}_{jmt}$ and $\widetilde{F}_{jmt}$ are the new equilibrium quantities and defaults under the counterfactual scenario. Following the non-monotonic price response predicted in the Monte Carlo experiment, we investigate what happens to equilibrium prices in this counterfactual scenario with respect to the actual prices. As shown in Figure 2.3, we find that almost all the prices vary, with the majority increasing by up to 5% (with some outliers not included in this figure), but some of them decreasing by at most 5%. We relate this price variation[38] to a measure of bank-province-year specific market power, which is the predicted markup derived from the demand

[38] We measure price variation as: $\Delta P_{jmt} = \frac{\widetilde{P}_{jmt} - P_{jmt}}{P_{jmt}}$.

model, i.e. the last term in equation (2.24).[39] We present this relationship in Figure 2.4, where we show evidence of a negative and statistically significant correlation between market power and price variation.[40] This means that prices increase in more competitive markets and decrease in more concentrated ones, confirming the predictions of the model.

We also look at the variation in quantities in the counterfactual scenario. We focus on the change in the share of borrowing firms[41] due to an increase in adverse selection. As shown in Figure 2.5, similarly to the price variation, we find that the share of borrowing firms varies in both direction, both increasing and decreasing. When we regress this share variation on the average markup at the province-year level, we find a positive and significant coefficient, presented in Figure 2.6. This suggests that the share of borrowing firms increases in more concentrated markets, as a natural consequence of the price reduction.

Last, we consider the effect of an increase in adverse selection on the average default rates[42] that a bank faces in a province-year. Figure 2.7 confirms that again most of the defaults are unchanged, but that a fraction of the banks in some provinces-years experiences either an increase or a reduction in borrowers' defaults. When we investigate the relationship between changes in default rates and markups, shown in Figure 2.8, we find that banks with a higher markup tend to reduce their default rates. This is the effect of the reduction in prices, which on one side attracts safer borrowers, and on the other reduces the probability of default of borrowers, given the estimated positive and significant effect of interest rates on defaults in Table 2.3.

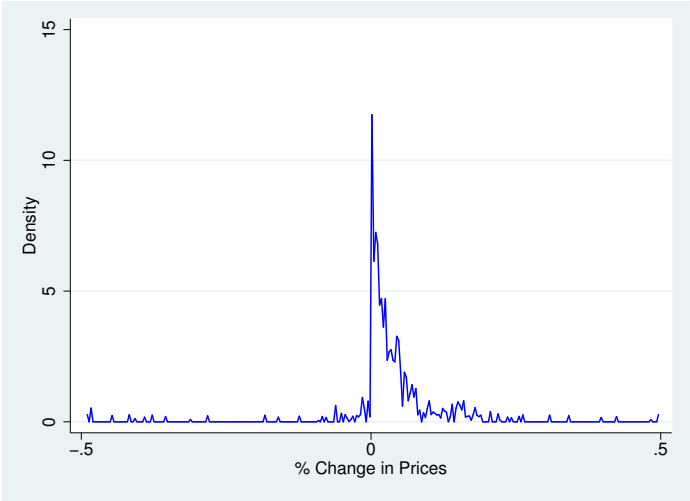
[39] We use the markup estimated in the baseline model. We tried also with other measures of competition at the local market level, like HHI of branches and loans, with similar results.

[40] We run a regression of bank-province-year level price variation on bank-province-year level markup, controlling for province-year fixed effects, and find a negative and significant coefficient.

[41] We measure this as the variation in the sum of the shares of the inside goods: $\Delta \sum_j S_{jmt} = \sum_j \widetilde{S}_{jmt} - \sum_j S_{jmt}$

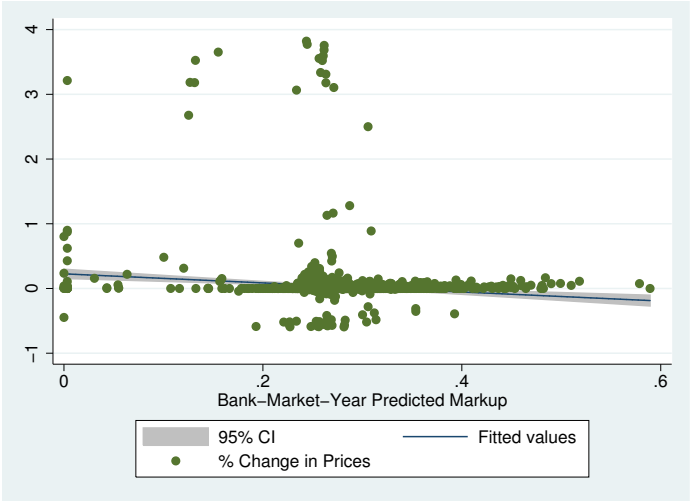
[42] We measure default rates' variation as: $\Delta F_{jmt} = \widetilde{F}_{jmt} - F_{jmt}$

Figure 2.3: Kernel Density of Price Variations



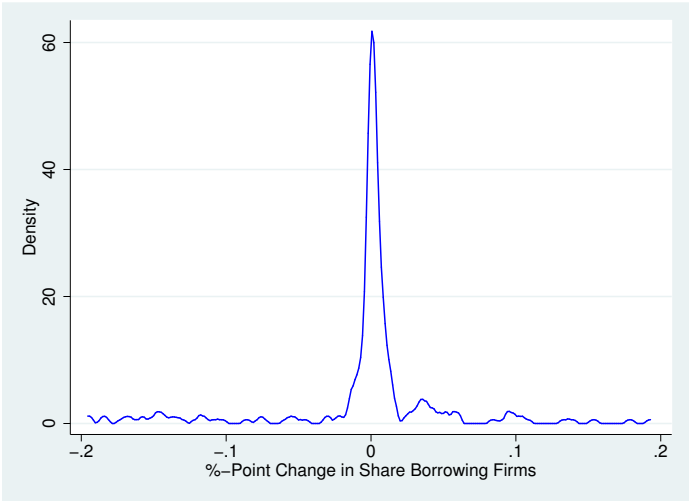
Note: The vertical axis is the density. The horizontal axis is the percentage variation between actual and counterfactual prices.

Figure 2.4: Regression of Price Variation on Markup



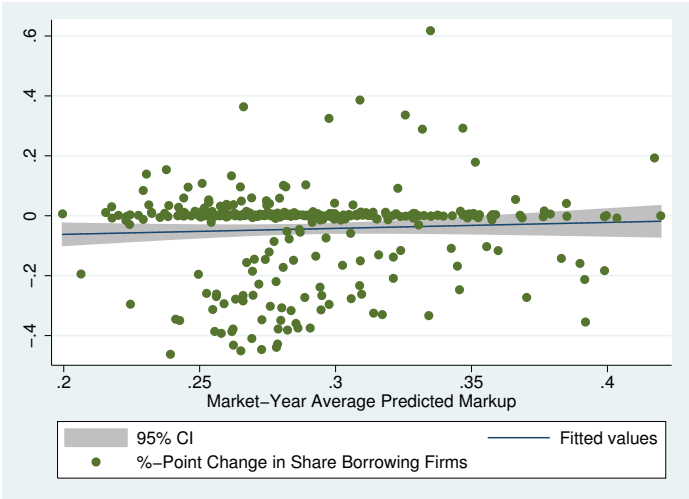
Note: The vertical axis is the percentage variation between actual and counterfactual prices. The horizontal axis is the bank-province-year level measure of markup.

Figure 2.5: Kernel Density of Quantity Variations



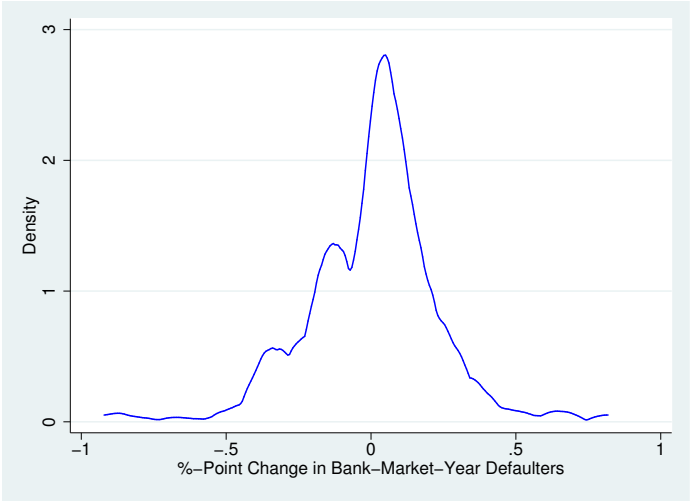
Note: The vertical axis is the density. The horizontal axis is the variation between actual and counterfactual share of the borrowing firms in each province-year in percentage points.

Figure 2.6: Regression of Quantity Variation on Markup



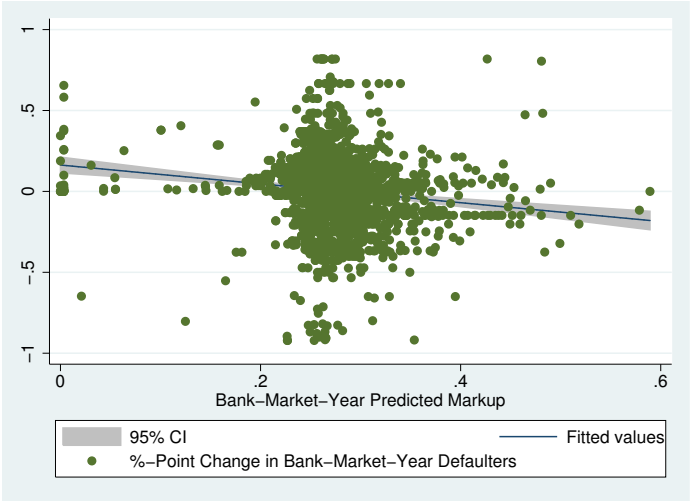
Note: The vertical axis is the variation between actual and counterfactual share of the borrowing firms in each province-year in percentage points. The horizontal axis is the percentage variation between actual and counterfactual prices.

Figure 2.7: Kernel Density of Default Variations



Note: The vertical axis is the density. The horizontal axis is the variation between actual and counterfactual default rates in percentage points.

Figure 2.8: Regression of Default Variation on Markup



Note: The vertical axis is the variation between actual and counterfactual default rates in percentage points. The horizontal axis is the bank-province-year level measure of markup.

2.8 Conclusion

In this paper we analyzed the interaction between imperfect competition and asymmetric information in the Italian market for small business lines of credit. We have access to a rich dataset with detailed information about credit line contracts between firms and banks, including all the main Italian credit institutions and a highly representative sample of firms. Using this data, we provide reduced form evidence of adverse selection in the spirit of the positive correlation test on unobservables by Chiappori and Salanié [2000]. We find stronger presence of asymmetric information for new borrowers.

Based on this evidence, we propose a structural model of firms' demand for credit, loan use, and default, as well as of banks' pricing. We let differentiated banks compete à la Bertrand-Nash on interest rates in local credit markets, but also use interest rates as a screening device, as in Stiglitz and Weiss [1981]. The model allows for imperfect competition in the lending market, accounting for asymmetric information between borrowers and lenders. We assume in fact that firms know the riskiness of their own project, but banks can only observe the average riskiness of their borrowers, conditional on observable firm characteristics. When we introduce asymmetric information, our model of oligopolistic competition predicts different banks' interest rate reactions, depending on the level of competition. We provide Monte Carlo evidence of a non-monotonic optimal bank's price response to an increase in adverse selection, depending on different measures of competition. More adverse selection causes prices to increase in competitive markets, but can have the opposite effect in more concentrated ones, where banks can leverage over their markup to lower prices and attract safer borrowers.

We find evidence of adverse selection in the data, both in the form of a positive correlation between unobservables determining demand and default, and unobservables affecting the size of the loan and default. We conduct a policy experiment to simulate the effects of a credit crunch, in which risky firms experience a more severe financial distress and demand more credit, doubling adverse selection. As predicted, in this counterfactual scenario equilibrium prices rise for more competitive markets and decline for more concentrated ones. Moreover, the share of borrowing firms increases in more concentrated markets, and default rates fall.

2.9 Appendix A - Constructing the Dataset

We have assembled various datasets from different sources, which are the following:

- **Firm Data:** Dataset from *Centrale dei Bilanci* with yearly (1988-1998) balance sheet data for each firm, including both firms that take credit and don't (outside option). This also includes the year of birth of each firm and its location at the city council level.
- **Score Data:** Dataset for each firm with yearly (1982-1998) score data, with also the 6 years preceding 1988. We retain from this data the 1982-1987 average, standard deviation, and weighted average (more weight to more recent years) of the score.
- **Loan Data:** Dataset from *Centrale dei Rischi* with yearly (1988-1998) firm-bank loan contracts, including amount granted, amount used, interest rate, firm's default. This is only for the main 94 banks and for short term credit lines.
- **Bank Data:** Dataset with yearly (1988-2002) balance sheet data for each bank, including yearly total loans that each bank gives in each province, and its share of the total loans granted in each province.
- **Branch Data:** Dataset with yearly (1959-2005) branches for each bank at the city council level. This includes the population of banks ($\sim$ 1,500 banks).
- **Coordinates Data:** Based on the *ISTAT* city council classification, we assign to each city council the geographic coordinates that will allow us to calculate firm-branch distances.

We first merge the firm and score datasets with the loan data, in order to have all the borrowing and not borrowing firms together. We then take all the banks actively lending in each province and assume that those represent the choice set for each firm, regardless of whether they have a branch in that province or not [43]. We assume that each firm chooses one main credit line among all the banks available in its province. The main line is defined as the line for which the amount used, regardless of the amount granted, is the highest. For cases in which multiple

[43] There is evidence in other papers (Bofondi and Gobbi [2006]), as well as in our data, that a few banks lend in some provinces even if they don't have a branch there.

lines have the same amount used, then the one with the lowest price is chosen. We calculate the distance in km between the city council of each firm and the city council where each bank from the choice set has a branch using the geographic coordinates. For each firm-bank pair, we only keep the branch that is closest to the firm.

2.9.1 Predicting Prices and Amounts Granted

We only consider the first year in which a firm appears in our sample. We assume that banks have a posted price for each observable type of firm in each market, defined as a year-province combination. We recover this synthetic price using regression analysis based on the actual prices that we observe.[44] We need to do this to predict the price that would have been offered to firms not borrowing in the data, as well as the price that would have been offered to borrowing firms by banks other than the chosen one. For this reason, we use not only the interest rate charged for the main credit line, but also the rates for the other lines that a firm opens in its first year. We run the following OLS regression:

$$P_{ikjmt} = \underbrace{\widehat{\alpha} + \widehat{\lambda}_{jmt}}_{\widetilde{P}_{jmt}} + \underbrace{\widehat{\omega}_k}_{\widetilde{\Delta P}_k} + \varepsilon_{ikjmt}, \tag{2.26}$$

where P_{ikjmt} is the interest rate that bank j charges to firm i of type k in market m at time t, λ_{jmt} is a bank-market-time interactive fixed effect and ω_k are interactive type dummies. We define an observable type based on the amount granted, the sector, the size in terms of sales, and the observable riskiness (SCORE). The underlying assumption here is that the effect on prices of the observable type's characteristics is additively separable with respect to the bank-market effect. Table 2.5 summarizes all the categories that define a type, as well as some interest rate descriptives for each value of these categories. This regression allows us to recover the bank-market-time specific average price $\widetilde{P}_{jmt}$, as well as the type specific deviation from this average $\widetilde{\Delta P}_k$. Given the large number of variables estimated, we don't report the results of the regression.[45] However, given that in the estimation we will use only the bank-market-time specific average prices, we present some descriptive statistics comparing the predicted prices $\widetilde{P}_{jmt}$ with the actual prices P_{jmt} in the data in Table 2.6, as well as two overlapping kernel densities in Figure 2.9, to show the goodness of fit of the model.

We are mainly interested in prices at the bank-market level because we use

[44] We are working on alternative ways of predicting prices, following Gerakos and Syverson [2014].

[45] We find that the bank-market-time fixed effects as well as the types' fixed effects are jointly significant, and the R^2 is 0.5092.

this variable only in the second stage of our estimation. We provide the intuition of our approach with a simplified version of the demand model. Let the utility of firm i of type k in market m at time t to choose bank j be:

$$U_{ikjmt} = \delta_{jmt} + \gamma_k + \varepsilon_{ikjmt}, \tag{2.27}$$

in the first stage of our estimation we recover $\widehat{\delta}_{jmt}$ and $\widehat{\gamma}_k$, that are bank-market fixed effects and types' fixed effects, which in the second stage are used to derive the price coefficients. We are just interested in the price coefficient at the bank-market-time level, and the price variation at the type level will be captured by γ_k, so we will only run the following second stage:

$$\widehat{\delta}_{jmt} = \alpha_0 + \alpha_1 \widetilde{P}_{jmt} + \beta X_{jmt} + \xi_{jmt}. \tag{2.28}$$

Similarly to the price, we also need to predict what the amount granted would be for firms that don't borrow. We do so using regression analysis from the borrowing firms. This is simplified by the fact that the distribution of amounts granted among the borrowing firms shows evident mass points corresponding to round numbers (mostly between 50 and 500 thousands euros), which are strongly correlated with several firm characteristics (for example, bigger firms get a greater amount). Given that we just need one amount granted for each non-borrowing firm, we calculate the median amount granted to each firm in our data, and group the resulting amounts in the 5 categories listed in Table 2.5. We regress them against several firm level controls[46] and a province-year-sector interactive fixed effect. The model predictions compared to the actual amounts are shown in Table 2.7. The model performs relatively well for the most demanded amounts (between 50 and 500 thousand euros), but performs poorly for the least demanded ones (below 50 and above 500 thousands).

[46] Tangible and intangible assets, total assets, net assets, short term debt, sales, profits, cashflow, SCORE, long term and short term total bank debt, returns on assets, age of the firm.

Table 2.5: Types' Summary Statistics

Category				Interest Rate		
		Obs	Percent	Mean	Median	Std. Dev.
Amount Granted	0 - 50,000	12,135	12.16	16.75	16.34	4.70
	50,001 - 100,000	17,014	17.05	15.60	15.19	4.24
	100,001 - 200,000	22,823	22.87	14.65	14.26	3.92
	200,001 - 500,000	27,440	27.50	13.81	13.39	3.67
	500,001 - 3,000,000	20,363	20.41	12.45	12.15	3.35
Sector	Primary	17,076	17.11	14.30	13.70	4.29
	Secondary	42,775	42.87	14.53	14.01	4.17
	Tertiary	39,924	40.01	14.27	13.73	4.01
Size	Small	40,223	40.31	15.45	14.95	3.93
	Large	59,552	59.69	13.67	13.07	4.11
SCORE	Low Risk	25,684	25.74	13.95	13.32	4.28
	Medium Risk	33,659	33.73	14.20	13.72	4.09
	High Risk	40,432	40.52	14.82	14.3	4.02
Total		99,775	100.00	14.39	13.82	4.13

Note: Interest rates is winsorized for the top and bottom 1% of its distribution. We exclude loans above 3,000,000 euros, which represent 2.5% of the loans in our sample. Primary sector includes primary, minerals' extraction, chemicals, metals, energy. Secondary sector includes food and beverages, textile and clothing, wood, paper and publishing, mechanical and electronic machines, production of transport vehicles, other manufacturing, and constructions. Tertiary sector includes commerce of transport vehicles, other commerce, hotels and restaurants, transport, storing and communications, real estate, financial intermediaries, and public administration. Size is defined as firms above or below the median of the distribution of yearly sales, which is around 10 million euros. Low risk is for SCORE values between 1 and 4, medium risk between 5 and 6, and high risk 7 to 9.

Table 2.6: Descriptives Comparing Actual and Predicted Prices

Statistics	**Actual Price**	**Predicted Price**
Mean	16.594	16.592
Standard Deviation	2.579	2.579
10th Percentile	13.400	13.399
50th Percentile	16.500	16.500
90th Percentile	19.954	19.954
Correlation Coefficient		1.000***
P-Value		0.000

Note: These are average prices at the year-province-bank level for the types' default categories: 0-50,000 euros, primary sector, small size, low risk.

Figure 2.9: Kernel Densities Comparing Actual and Predicted Prices

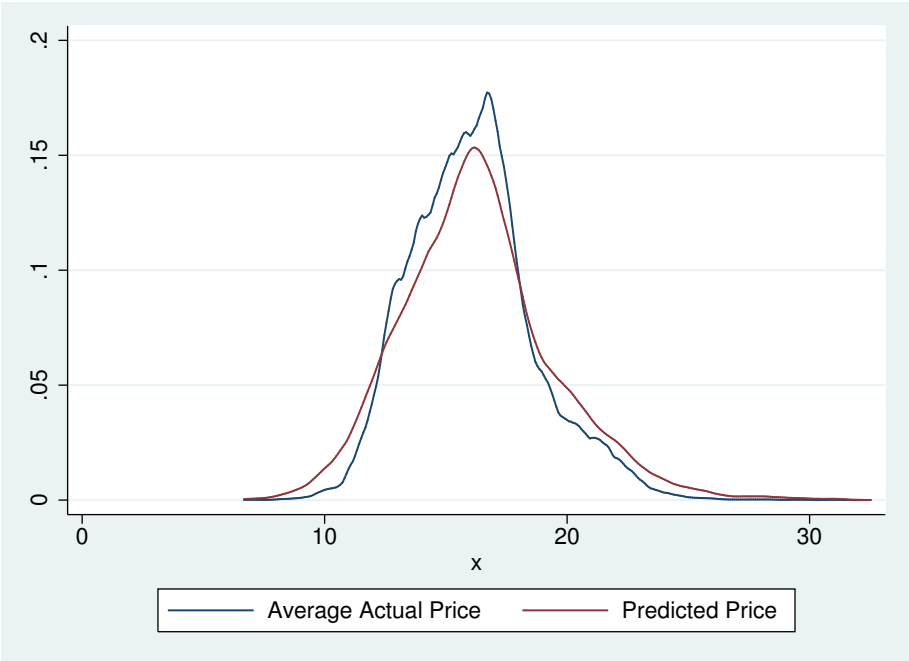


Table 2.7: Percentage of Predictions and Actual Amounts Granted in Thousands of Euros

	Actual					
Predicted	0 - 50	51 - 100	101 - 200	201 - 500	501 - 3,000	**Total**
0 - 50	2.6	1.7	0.4	0.1	0.0	0.72
51 - 100	23.7	18.6	10.9	4.5	1.2	10.19
101 - 200	63.3	65.8	66.6	59.6	40.0	59.90
201 - 500	10.4	13.9	22.1	35.8	58.2	29.08
501 - 3,000	0.0	0.0	0.0	0.0	0.6	0.11
Total	5,667	11,222	15,580	17,301	9,268	59,038

Note: Each column sums up to 100%. The last column on the right represents the predicted total number of observations for each mass point, whereas the last row represents the actual total number of observations for each mass point.

2.10 Appendix B - IV First Stage and OLS vs IV Second Stage

Table 2.8: IV First Stage and OLS vs IV Second Stage for Demand

	First Stage	**Second Stage**	
Variable	**Interest Rate**	**OLS**	**IV**
Share of Branches of Merging Rivals	0.215***	-	-
	(0.021)		
Interest Rate	-	4.506***	-3.669***
		(0.303)	(0.348)
Number of Branches	-0.014	-2.766***	-2.746***
	(0.016)	(0.233)	(0.269)
Share of Branches	0.200***	11.497***	12.646***
	(0.043)	(0.627)	(0.721)
Years in Market	-0.030***	-0.799***	-1.001***
	(0.010)	(0.150)	(0.172)
Constant	0.801***	-2.111***	4.580***
	(0.012)	(0.306)	(0.352)
Bank FE	Yes	Yes	Yes
Obs	2,279	2,279	2,279
R^2	0.2028	0.2943	-
F-Stat	97.606	-	-

Note: Standard errors in brackets. $*$ is significant at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

Table 2.9: IV First Stage and OLS vs IV Second Stage for Loan Size

	First Stage	**Second Stage**	
Variable	**Interest Rate**	**OLS**	**IV**
Share of Branches of Merging Rivals	0.215***	-	-
	(0.021)		
Interest Rate	-	0.030	-0.295***
		(0.097)	(0.097)
Number of Branches	-0.014	-0.103	-0.102
	(0.016)	(0.075)	(0.075)
Share of Branches	0.200***	0.383*	0.429**
	(0.043)	(0.200)	(0.201)
Years in Market	-0.030***	-0.024	-0.032
	(0.010)	(0.048)	(0.048)
Constant	0.801***	0.070	0.337***
	(0.012)	(0.098)	(0.098)
Bank FE	Yes	Yes	Yes
Obs	2,279	2,279	2,279
R^2	0.2028	0.0425	-
F-Stat	97.606	-	-

Note: Standard errors in brackets. $*$ is significant at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

Table 2.10: IV First Stage and OLS vs IV Second Stage for Default

	First Stage	**Second Stage**	
Variable	**Interest Rate**	**OLS**	**IV**
Share of Branches of Merging Rivals	0.215***	-	-
	(0.021)		
Interest Rate	-	-0.283	2.387***
		(0.384)	(0.389)
Number of Branches	-0.014	-0.230	-0.236
	(0.016)	(0.297)	(0.300)
Share of Branches	0.200***	0.473	0.098
	(0.043)	(0.796)	(0.805)
Years in Market	-0.030***	-0.190	-0.124
	(0.010)	(0.191)	(0.193)
Constant	0.801***	-0.727*	-2.913***
	(0.012)	(0.389)	(0.393)
Bank FE	Yes	Yes	Yes
Obs	2,279	2,279	2,279
R^2	0.2028	0.1740	-
F-Stat	97.606	-	-

Note: Standard errors in brackets. $*$ is significant at the 10% level, $**$ at the 5% level, and $***$ at the 1% level.

Chapter 3

Dynamic Entry and Exit with Learning by Branching

Dynamic Entry and Exit with Learning by Branching*

Nicola Pavanini†, Fabiano Schivardi‡

June 2014

*We thank Lorenzo Ciari, Greg Crawford, Michael McMahon, Carlos Noton, Pasquale Schiraldi, Michael Waterson and seminar participants at the Warwick Economics PhD Conference 2013 and EARIE 2013 for helpful comments. We thank for financial support the Research Centre Competitive Advantage in the Global Economy (CAGE), based in the Economics Department at University of Warwick.

†University of Zürich, nicola.pavanini@econ.uzh.ch

‡LUISS, EIEF and CEPR, fschivardi@luiss.it

Abstract

Asymmetric information is a potential determinant of market structure in the banking industry. When incumbents have an informational advantage over potential entrants, entry can be less profitable and concentration may arise. Incumbent banks can gain an informational advantage in a market by accumulating experience over time, through learning about the market and building relationships with the borrowers. However, the cost of learning depends on the quality of the pool of borrowers that the bank lends to. This paper wants to quantify the cost of asymmetric information as an entry barrier in the Italian market of credit lines to small businesses. We do so measuring welfare under different counterfactual scenarios, with and without borrower-lender asymmetric information and reducing incumbents' informational advantage. First, we provide reduced form evidence of learning by branching, showing that banks improve the quality of their borrowers as they accumulate experience over the years. We also show that entry is negatively correlated with concentration of experience. Given this, we estimate a dynamic structural game of banks' entry, exit and investment with learning by branching based on Weintraub et al. [2008b], together with a static framework of firms' demand for credit, loan size, default and banks' pricing, that allows us to identify the effect of asymmetric information on market structure.

3.1 Introduction

Asymmetric information is a potential determinant of market structure in the banking industry. When incumbents have an informational advantage over potential entrants, entry can be less profitable and concentration may arise. Incumbent banks can gain an informational advantage in a market by accumulating experience over time, through learning about the market and building relationships with the borrowers. However, the cost of learning depends on the quality of the pool of borrowers that the bank lends to. For a new entrant, the creditworthiness of its borrowers will be influenced by the informational advantage that the incumbent banks have. This paper wants to determine the cost of asymmetric information as an entry barrier in the Italian market of credit lines to small businesses. We do so measuring welfare under different counterfactual scenarios, with and without borrower-lender asymmetric information and reducing incumbents' informational advantage.

Several papers in the banking literature analyse the effects of asymmetric information on competition and market structure. In the context of incumbent banks more informed than potential entrants, the equilibrium results found are blockaded entry (Dell'Ariccia et al. [1999]), a finite number of banks (Dell'Ariccia [2000], Dell'Ariccia [2001]), and larger adverse selection cost for entrants (Marquez [2002]). Petersen and Rajan [1995] argue that credit market competition with asymmetric information may be inimical to the formation of relationships between firms and banks. Pagano and Jappelli [1993] show that lenders' incentives to share information about borrowers are reduced by the fear of competition by potential entrants. Broecker [1990] finds an oligopoly result when banks compete on interest rates using imperfect and independent test of borrowers' creditworthiness. The empirical literature provides some evidence of higher loan default rates for new entrants, using data of the US (Shaffer [1998]) and Italy (Bofondi and Gobbi [2006]). Additionally, Gobbi and Lotti [2004] show that incumbents have an informational advantage over new entrants in the Italian credit market.

An intuitive way for lenders to overcome informational asymmetries and the inefficiencies these might cause is to learn about their borrowers. However, learning also generates an informational advantage for experienced incumbent lenders versus potential entrants. Following this direction, some of the theoretical papers in this strand of literature introduce the concept of learning by lending, which can justify the endogenous creation of an informational advantage for incumbents (Dell'Ariccia [2000], Dell'Ariccia [2001]). This idea is borrowed from an earlier literature on the strategic implications of learning by doing. Based on the seminal papers by Arrow

[1962] and Spence [1981], Dasgupta and Stiglitz [1988] find that learning acts as an entry barrier, whereas Cabral and Riordan [1994] show that learning leads to increasing dominance by incumbents. More recently, Besanko et al. [2010] developed a general model of dynamic competition based on Ericson and Pakes [1995] that accounts for learning by doing and organisational forgetting, and find that these forces can deliver varying degrees of long-run industry concentration. There is also a recent growing interest in structural IO models that allow to recover learning and forgetting parameters (Benkard [2000], Benkard [2004]).

In this paper we try to capture these effects estimating a dynamic structural model of banks' entry, exit and investment through branching. Together with a companion framework estimating a static model of firms' demand for credit, loan size, default, and banks' pricing (Crawford et al. [2013]), we identify the effect of asymmetric information on market structure. Given the large number of banks in each market, and the nonstationary environment that we face after branching deregulation, we adopt the nonstationary oblivious equilibrium notion developed by Weintraub et al. [2008b] and Weintraub et al. [2008a]. We estimate this model using a unique set of linked datasets of the Italian market for small business lines of credit. We consider the period 1988-1998, right after entry deregulation, during which there were substantial changes in concentration and market structure of local credit markets. We provide reduced form evidence of learning by branching, showing that banks improve the quality of their borrowers as they accumulate experience over the years. We also show that entry is negatively correlated with concentration of incumbents' experience, and that exit is negatively correlated with a bank's own experience level.

The contribution of this paper is twofold. On one hand, it contributes to the structural IO and empirical banking literature developing a dynamic game of banks' entry and exit that, differently from existing papers (de Elejalde [2012]), allows for pricing competition and investment in branching. On the other, it provides a new empirical insight on the effects of asymmetric information on market structure, that so far have mostly been investigated from a theoretical and reduced form perspective. The paper is organized as follows. We present the data and the institutional background in Section 2, descriptives and reduced form evidence of learning by branching, entry and exit in Section 3, the structural model in Section 4, the estimation methodology in Section 5, counterfactuals in Section 6, and then the conclusions.

3.2 Data and Institutional Details

We have access to a unique dataset of small business credit lines, previously used in Panetta et al. [2009]. We use three main sources of data. Interest rate data and data on outstanding loans are from the Italian *Centrale dei Rischi*, or Central Credit Register. Firm-level balance sheet data are from the *Centrale dei Bilanci* database. Banks' balance-sheet and income-statement data are from the Banking Supervision Register at the Bank of Italy. By combining these data, we obtain a matched panel dataset of borrowers and lenders extending over an eleven-year period, between 1988 and 1998.

The Central Credit Register (hereafter CR) is a database that contains detailed information on all individual bank loans extended by Italian banks. Banks must report data at the individual borrower level on the amount granted and effectively utilized for all loans exceeding a given threshold [1], with a breakdown by type of the loan (credit lines, financial and commercial paper, collateralized loans, medium and long-term loans and personal guarantees). In addition, a subgroup of around 90 banks (accounting for more than 80 percent of total bank lending) have agreed to file detailed information on the interest rates they charge to individual borrowers on each type of loan. Summary statistics for these banks are reported in Panel A of Table 3.1.

We restrict our attention to short-term credit lines, which have ideal features for our analysis. First, the bank can change the interest rate at any time, while the borrower can close the credit line without notice. This means that differences between the interest rates on loans are not influenced by differences in the maturity of the loan. Second, the loan contracts included in the CR are homogeneous products (for example, they are not collateralized), so that they can be meaningfully compared across banks and firms. Third, short term bank loans are the main source of borrowing of Italian firms. For example, in 1994 they represented 53 percent of the total debts according to the Flow of Funds data. We define the interest rate as the ratio of the payment made in each year by the firm to the bank to the average amount of the loan. The interest payment includes the fixed expenses charged by the bank to the firm (e.g. which encompass the cost of opening the credit line or the cost of mailing the loan statement).

The *Centrale dei Bilanci* (hereafter CB) collects yearly data on the balance sheets and income statements of a sample of about 35,000 Italian non-financial and non-agricultural firms. This information is collected and standardized by a con-

[1] The threshold was 41,000 euros (U.S. $42,000) until December 1995 and 75,000 euros thereafter.

sortium of banks interested in pooling information about their customers. A firm is included in the CB sample if it borrows from at least one of the banks in the consortium. The database is fairly representative of the Italian non-financial sector. The firms in the CB sample represent about 49.4% of the total sales reported in the national accounting data for the Italian non-financial, non-agricultural sector. Table 3.1, Panel B reports descriptive statistics for the sample of borrowing and non-borrowing firms. These two groups of firms appear to be fairly similar in terms of size, leverage and riskiness, but as expected borrowing firms have a higher share of short term debt compared to non-borrowing ones. The unique feature of the CB data set is that, unlike other widely used data sets on individual companies (such as the Compustat database of US companies), it has wide coverage of small and medium companies; moreover, almost all the companies in the CB sample are unlisted. The coverage of these small firms makes the data set particularly well suited for our analysis, because informational asymmetries are potentially strongest for these firms.

In addition to collecting the data, the CB computes an indicator of the risk profile of each firm (which we refer to in the remainder of this paper as the SCORE). The SCORE represents our measure of a firm's observable default risk. It takes values from 1 to 9 and is computed annually using discriminant analysis based on a series of balance sheet indicators (assets, rate of return, debts etc.) according to the methodology described in Altman [1968] and Altman et al. [1994].

3.2.1 Local Credit Markets

Following other papers on Italian local credit markets (Felici and Pagnini [2008], Bofondi and Gobbi [2006], Gobbi and Lotti [2004]), we identify banking markets as the Italian provinces. As for many other industries, there is no clear consensus on what should be the most appropriate market definition. What is known from the seminal work of Bresnahan and Reiss [1991] is that local markets should be independent and isolated. This is explained in more detail for the banking sector by Cohen and Mazzeo [2007], who claim that in a local banking market consumers should not typically use depository institutions outside of their area, and that distinct or overlapping submarkets should not exist within the defined geographic markets. Claiming that Italian provinces are independent and isolated markets according to the definition of Cohen and Mazzeo [2007] is a strong assumption. However, limiting the analysis to small isolated markets would substantially hamper the welfare and policy implications of our analysis. Moreover, provinces are also used by Italian

supervisory authorities as proxies for the local markets for deposits.[2]

We define entry as the act of starting to lend in a geographic market, regardless of whether the bank has an established branch or not, and define exit similarly. Table 3.2 shows summary statistics on local credit markets. On average, there are 2 new banks entering every period in every market, and 1.6 banks exiting. The substantial turnaround that we observe in the data during these years is the effect of the entry deregulation at the end of the 1980s, after a long period of geographic and scope restrictions for Italian banks, as documented in Guiso et al. [2007]. In particular, right before deregulation in 1985 Italy had only 0.23 branches per 1,000 inhabitants, whereas the EU average was more than double (0.52 per 1,000 inhabitants), despite the Italian economy being very close to the main European economies in terms of GDP per capita in those years.[3] Moreover, by 1995 Italy experienced a 78.3% growth rate in the number of branches per 1,000 inhabitants, compared to the EU average growth rate of -5.6%.[4]

According to the definition of the US department of Justice [5], on average these markets are moderately concentrated, with the top 10% being highly concentrated. On average, 2.2% of borrowers default. A relationship between a borrower and a lender within the sample last for around 2.6 years.

[2] Ciari and Pavanini [2013], who look at market structure and multi market contact in the Italian banking sector, have used the notion of Local Labor Systems as relevant markets. These are geographical zones defined according to integrated economic areas and commuting patterns between councils. We are working on experimenting with this alternative definition to compare the sensitivity of our results to the choice of market borders.

[3] The main European countries' branches per 1,000 inhabitants and GDP per capita (in 2008 US$), from ECB and World Development Indicators (2008) data, are: France 0.47 (9,823 US$), Germany 0.61 (9,125$), UK 0.38 (8,062$), Italy 0.23 (7,699$), Spain 0.76 (4,569$).

[4] The main European countries' growth rates in terms of branches per 1,000 inhabitants were: France -6.4%, Germany -3.3%, UK -14.2%, Italy 78.3%, Spain 22.4%.

[5] Based on the definition of the "Horizontal Merger Guidelines" issued by the U.S. Department of Justice and the Federal Trade Commission on August 19th, 2010.

Table 3.1: Summary statistics: Banks and Firms

Variable	Obs.	Mean	Stand. Dev.	5th pctile	Median	95th pctile
			Panel A: The Bank Sample			
Total Assets	900	10,726.8	16,965.6	481.3	3,709	54,354.1
Employees	896	3,179.9	4,582.5	206	1,137	14,038
Bad Loans	893	6.2	6.3	1.9	4.9	15.8
Costs-Income ratio	893	34.5	6.1	25.4	33.1	43.2
			Panel B.1: The Borrowing Firm Sample			
Total Assets	302,747	8.1	12.2	0.9	4.2	29.1
Employees	272,816	54.4	75.6	3	30	195
Leverage	305,151	0.57	0.28	0	0.62	0.95
Return on Sales	301,821	1.1	7.5	-9.7	1.2	11.1
Short Term Debt	305,752	33	22.9	0	32	70.7
SCORE	307,532	5.2	1.8	2	5	8
No. of Lenders	329,623	4.4	3.3	1	4	11
Utilized Credit	319,792	50.2	54.3	0	38.2	138.4
			Panel B.2: The Non-Borrowing Firm Sample			
Total Assets	209,754	8.8	20.1	0	2.8	38.8
Employees	176,248	60.6	124.4	0	20	269
Leverage	208,441	0.49	0.36	0	0.52	1
Return on Sales	197,624	2.2	19.6	-17.1	1.2	22.4
Short Term Debt	195,663	23.9	25.7	0	15.7	73.9
SCORE	206,378	4.8	2.1	1	5	8

Note: An observation is the number of bank-years with non-missing records in Panel A, and firm-years in Panel B. Total assets are in millions of euros. Employees is the number of employees at the end of the year. Bad loans is a percentage of total loans. Cost-income ratio is the ratio of overhead to gross income (in %). Return on sales is calculated as the percentage ratio of current profits over total sales. Short term debt is expressed as a proportion of total debt. The SCORE is the indicator of the risk of the company computed each year by the Centrale dei Bilanci (higher values indicate riskier companies). Number of lenders is the number of banks from which the company borrows. Utilized credit is expressed as a proportion of credit granted. The first five variables in each of the firm's sample are winsorised at the 1st and 99th percentile.

Table 3.2: Summary statistics: Local Credit Markets

Variable	Obs.	Mean	Stand. Dev.	10th pctile	Median	90th pctile
Yearly Amount Lent	1,133	234.7	568.1	22.4	91.8	459.7
Average Interest Rate	1,133	14.78	2.35	11.43	14.88	17.04
N of Banks	1,133	26.65	12.14	12	25	43
N of Entering Banks	1,030	1.96	1.77	0	2	4
N of Exiting Banks	1,030	1.6	1.45	0	1	3
HHI	1,133	1,537.31	695.09	855.77	1,366.97	2,479.49
N of Firms	1,133	277.06	479.3	35	135	655
% of Firms Defaulting	1,133	2.2	2.8	0	1.2	6.1
Avg. Rel. Length	103	2.63	0.17	2.44	2.64	2.83

Note: An observation is a province-year combination for all the variables but the last one, where an observation si a province. Amount lent is in million of euros.

3.3 Descriptives and Reduced Form Evidence

3.3.1 Descriptive Statistics

Before introducing some regression results and outlying the features of the structural model, we want to give a sense of the variation in the data that motivates our analysis. We are investigating whether asymmetric information between borrowers and lenders can serve as an entry barrier in the Italian banking sector. This occurs when incumbent banks have been learning about their borrowers through years of experience in a market, and are able exploit their informational advantage against new banks that enter a market without any knowledge of borrowers' creditworthiness. We provide preliminary supporting evidence of this effect, showing that incumbents maintain a large share of the market, and have on average a safer pool of borrowers, to which they are able to charge a higher rate compared to newly entering banks.

We focus on a subsample of the data, comparing over a five years period the performance of banks starting to lend in a market for the first time to the existing incumbents. We take the years between 1990 and 1994, and define as incumbents those banks that were present in the market before 1988, and as entrants only the banks that started lending in a market for the first time in 1990.[6] We drop all the entrants after 1990 and simply compare some outcomes between incumbents and entrants over that period.

We start with Figure 3.1a, where we present the evolution of the average market share of incumbents and entrants. The figure shows how incumbents maintain a larger share of the market, and entrants slowly increase their quota over the years. Note however that on average entrants gain only about 5% market shares over this 5 years period. We also present evidence of how much of the credit granted borrowers end up using, summarized in Figure 3.1b. The graph shows that, at least for their first two years in a market, entrants attract borrowers that use a higher share of their loan with respect to the incumbents' borrowers. This suggests that the new entrants' first batch of borrowers is less capable of self-financing its projects.

One of the ways entrants try to increase their presence in the market is charging a lower rate than the incumbents, as shown in Figure 3.2a. Note that this interest rate difference between incumbents and entrants tends to persist over the years. We also take all the new entrants since 1990 and average their rates over the numbers of years they have spent in the market, in Figure 3.2b. This shows how

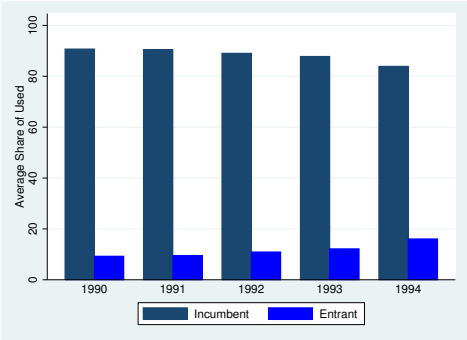
[6] We take this time period just as an example, and limit ourself to 5 years as in most of the descriptives the differences between incumbents and entrants either reduce substantially after that or show a clear enough trend.

entrants increase their rates as they gain more experience in a market.

Last, we present evidence of how new entrants attract a worse pool of borrowers with respect to incumbents. We distinguish between bank's observable and unobservable riskiness of its borrowers. First, we show in Figure 3.3a that the average borrowers' SCORE is higher for entrants, and takes about 3 years to reach the incumbents' level. Second, we present in Figure 3.3b the average default rates of banks' borrowers. Here there seems to be a strong difference between entrants and incumbents especially in the first year. Moreover, entrants' borrowers seem to be less creditworthy in years of crisis, like 1994, when the overall average default rate of firms peaked with respect to the previous years in our sample. We will investigate this effect in more detail in the next section.

Figure 3.1: Credit for Incumbent vs. Entrant Banks

(a) Market Shares



(b) Amount Used/Granted

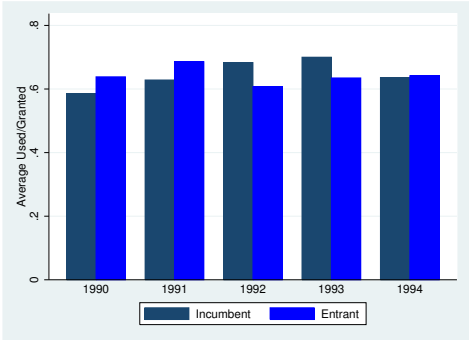
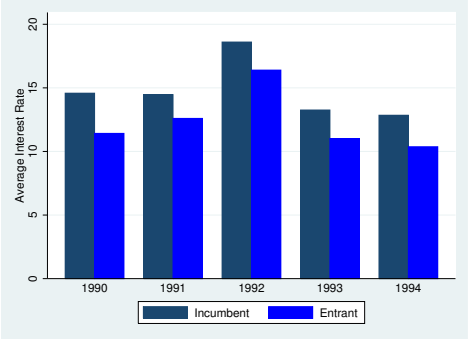


Figure 3.2: Interes Rates

(a) Incumbents vs Entrants



(b) Entrants over Time

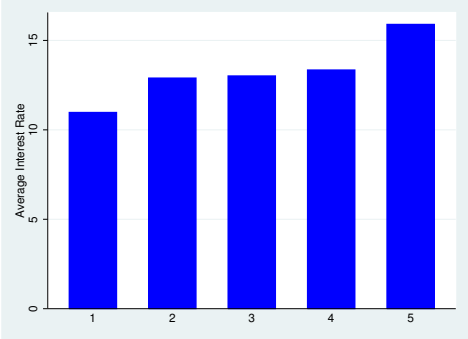
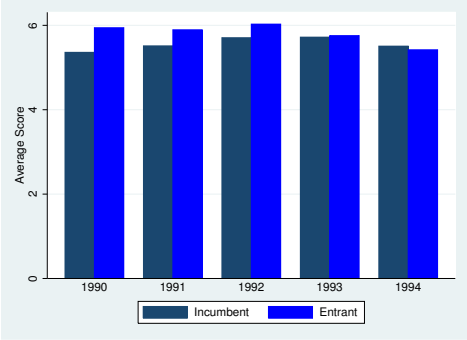
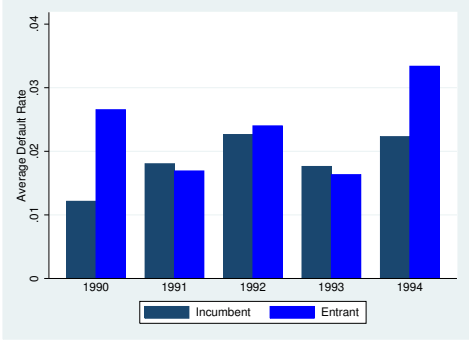


Figure 3.3: Borrowers' Riskiness for Incumbent vs. Entrant Banks

(a) Average Borrowers' SCORE



(b) Average Borrowers' Default Rates



3.3.2 Experience and Default

In this section we provide some reduced form evidence of the impact of banks' experience on default rates of its borrowers. We define experience as an unobserved stock of knowledge about a market and its borrowers that a lender accumulates over the years, which might also be regarded as the bank's stock of soft information. We assume that we can proxy experience with a function of the number of branches that a bank has in a market and the years that it has been there for. The idea is that the more branches a bank has, or the more years a bank is present for, the more it acquires information about a market, as it can screen and lend to a larger number of borrowers both in different parts of the market and across time. Thus, we investigate whether the cumulative share of branches in a market overtime improves the quality of the pool of borrowers in the following periods. The standard literature on learning by doing, as explained by Benkard [2000], would measure bank-market-year experience based on the following formula:

$$e_{jmt} = \delta e_{jmt-1} + q_{jmt-1}, \tag{3.1}$$

where the bank's stock of experience depreciates yearly by a factor $1-\delta$, also defined as organizational forgetting, and q_{jmt} is bank j's share of branches in market m at time t. We will use this baseline formula as a starting point for the structural model, but simplify it for the reduced form regressions, where we don't estimate δ, but assume experience is determined by a weighted average based on the harmonic series of past and current market shares. In particular, we define it as:

$$e_{jmt} = \sum_{\tau=1}^{t} \left(q_{jm\tau}\right) \frac{\tau}{t}. \tag{3.2}$$

This harmonic series allows us to give more weight to more recent market shares.[7] We run an OLS regression of the following baseline model:

$$d_{jmt} = \gamma e_{jmt} + \beta X_{jmt} + \delta_t + \omega_j + \kappa_m + \varepsilon_{jmt}, \tag{3.3}$$

where d_{jmt} is the default rate of the borrowers that bank j lends to in market m at year t. This rate is constructed as the ratio of number of defaulting firms over total number of firms the bank lends to in that market in that year. We control for year, bank and market fixed effects. We also control for other bank-year-market level variables, as the average interest rate, score and leverage of the bank's pool

[7] We have experimented with other definitions of experience, for example with cumulative amount lent, obtaining very similar results.

of borrowers. Residuals are clustered at the bank-market level. OLS estimates are summarized in Table 3.3, distinguishing between new borrowers, old borrowers, and the whole sample. We define new firms as firms that have never borrowed before from one lender (at least in our data), and old firms as borrowers that have at least one year of lending relationship with a bank.

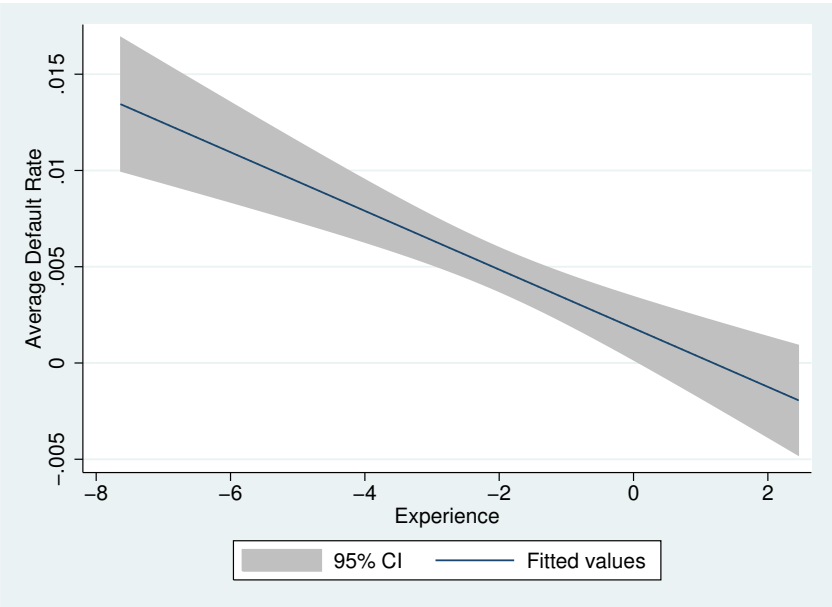
These results show that bank's experience is negatively and significantly correlated with the average default rate of both their old and new borrowers. The effect of experience is stronger for old firms. The estimates in Table 3.3 show that 1% increase in experience reduces average borrowers' default rate by 0.1 and 0.2 percentage points for new and old firms. Note that the mean of default rate is 6.3% and the standard deviation is 11%. The average rate, score, and leverage of borrowers have a positive and significant effect on default rates, as expected. In Figure 3.4 we report the correlation between default rates and experience, controlling for all the fixed effects and bank controls listed above. These results suggest that experience, measured as a function of number of branches and years in a market, is beneficial for the lender, as it improves its pool of borrowers overtime. This can be regarded as a valid motivation for developing a structural model that is able to explain the mechanism driving this result as well as its consequences in terms of market structure.

Table 3.3: Banks' borrowers' default rate

Variable	New Firms	Old Firms	All Firms
Experience	-0.001***	-0.002***	-0.002***
	(0.000)	(0.001)	(0.001)
Avg Rate	0.028***	0.037***	0.029***
	(0.005)	(0.009)	(0.007)
Avg Score	0.049***	0.118***	0.125***
	(0.006)	(0.016)	(0.017)
Avg Leverage	0.017***	0.030***	0.035***
	(0.004)	(0.007)	(0.006)
Obs	12,642	12,161	14,788
R^2	0.141	0.477	0.545
Year FE	Yes	Yes	Yes
Banks FE	Yes	Yes	Yes
Markets FE	Yes	Yes	Yes

Note: One observation is bank-year-market. Standard errors are clustered at the bank-market level. All variables are expressed in logs.

Figure 3.4: Average Borrowers' Default Rate and Banks' Experience



3.3.3 Experience and Entry and Exit

We also look at the relationship between entry, exit and experience in Table 3.4. Entry is defined as the act of lending for the first time in a market, and exit means stop lending in that market.[8] Following the example of the entry and exit policy functions estimated by Ryan [2012], we construct own e_{jmt} and rivals' e_{-jmt} experience variables, based on the definition described in the previous section. We also look at the effect of concentration of incumbents' experience ($HHIexp$), in terms of Herfindahl Hirschmann Index. We control for year, bank and market fixed effects. Residuals are clustered at the bank-market level. We run the following two probit models, whose marginal effects are reported in Table 3.4:

$$\begin{aligned}
\Pr(\chi_{jmt} = 1; e_{jmt} = 0) &= \Phi\big(\gamma_1 e_{-jmt} + \gamma_2 HHIexp_{mt} + \beta X_{jmt} + \delta_t + \omega_j + \kappa_m\big) \\
\Pr(\chi_{jmt} = 0; e_{jmt} > 0) &= \Phi\big(\gamma_1 e_{-jmt} + \gamma_2 HHIexp_{mt} + \gamma_3 e_{jmt} + \beta X_{jmt} + \delta_t + \omega_j + \kappa_m\big),
\end{aligned} \tag{3.4}$$

where $\chi_{jmt} = 1$ means entry and $\chi_{jmt} = 0$ means exit. We find that incumbents' experience has no effect on entry, but concentration of incumbents' experience has a negative and significant impact. This effect is present across different specifications. One unit increase in experience concentration (with mean 0.35, std dev 0.17) reduces entry probability by 0.037. We interpret this as additional potential evidence of entry barriers, as new entrants are less likely to approach markets where experience is concentrated among few incumbent banks.[9] We also show that banks' own experience has a negative and significant effect on exit. One unit increase in own experience (with mean 0.04, std dev 0.23) reduces exit probability by 0.052. The effect of own experience on exit is robust to various specifications. This result is complementary to what we find in the entry model, reflecting the fact that incumbents with more experience are less likely to exit a market, which implies that experience fosters their profitability.

[8] In this analysis we don't distinguish between greenfield entry (exit) and entry (exit) by merger and acquisition. However, most of the mergers and acquisitions in the Italian banking sector occurred in the beginning of the 2000s. See Ciari and Pavanini [2013] for a more extensive discussion on this point.

[9] A similar result has been found by Ciari and De Bonis [2011], who use an instrumental variables approach to show that Italian banks during the early stage of deregulation were more likely to open branches in more competitive markets rather than concentrated ones. They define competition as the difference between interest rates on loans and deposits, instrumented using pre-regulation (1936) market structure, as in Guiso et al. [2004].

Table 3.4: Banks' entry and exit decisions

Variable	**Entry**				**Exit**			
	(1)	**(2)**	**(3)**	**(4)**	**(1)**	**(2)**	**(3)**	**(4)**
Own Experience	-	-	-	-	-0.069***	-0.071***	-0.075***	-0.052***
					(0.011)	(0.011)	(0.011)	(0.006)
Rivals' Experience	0.022*	0.020	0.066	0.059	0.002	-0.007	-0.007	0.077
	(0.012)	(0.012)	(0.054)	(0.054)	(0.038)	(0.037)	(0.140)	(0.140)
HHI Experience	-0.028***	-0.028***	-0.037***	-0.037***	0.019*	0.020**	-0.071**	-0.050
	(0.003)	(0.004)	(0.013)	(0.013)	(0.010)	(0.010)	(0.036)	(0.033)
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Market FE	No	No	Yes	Yes	No	No	Yes	Yes
Banks FE	No	No	No	Yes	No	No	No	Yes
Obs	71,750	71,750	71,750	70,737	23,240	23,240	23,240	21,689
$PseudoR^2$	0.0036	0.0179	0.0352	0.1519	0.0135	0.0329	0.0459	0.3865

Note: One observation is bank-year-market. For entry, we allow each bank that is not already present to be a potential entrant in each market. For exit, we let each incumbent to potentially exit from each market where it's present. Standard errors are clustered at the bank-market level. All variables are in levels.

3.4 The Model

We build a structural model of firms' demand, loan size and default with asymmetric information, and of banks' entry, exit, investment and pricing. Given the facts presented in the reduced form evidence, we construct a model that allows banks to learn about their borrowers through experience, reducing defaults and increasing their profits. However, as pointed out in Besanko et al. [2010], learning can be used strategically by banks to promote their market dominance and generate concentration. As documented in Guiso et al. [2007], the strict entry regulation imposed on the Italian banking sector between 1936 and 1990 generated the unintended consequence of different degrees of competition across Italian provinces, and therefore different degrees of concentration. This means that at deregulation incumbent banks in highly concentrated markets had accumulated a higher level of experience compared to incumbents in low concentration markets, which granted them a bigger informational advantage with respect to new entrants. We investigate whether the higher information gap with incumbents represents a barrier to entry for new banks, as it makes learning for new entrants more costly in more concentrated markets. This resembles the theoretical results presented in Dell'Ariccia [2001], who shows that entry will be more difficult in markets where the institutional framework allows incumbent banks to acquire pervasive information about their clients.

3.4.1 Demand and Pricing

The demand and pricing side of the model is based on Crawford et al. [2013]. Assume there are $i = 1, ..., I$ firms of type $k = 1, ..., K$ and $j = 1, ..., J$ banks in $m = 1, ..., M$ markets. We omit the k subscript for simplicity. Let firms have the following utility from credit, which determines their demand:

$$U^D_{ijm} = \alpha^D_i + \delta^D_{jm}(X_{jm}, P_{jm}, \xi^D_{jm}, \beta^D) + V^D_{ijm}(Y_{ijm}, \eta^D) + \varepsilon^D_{ijm}. \tag{3.5}$$

We normalize to zero the utility from the outside option, which is not borrowing. Firms will choose the bank that maximizes their utility, or will choose not to borrow. Then, conditional on borrowing, they will choose the share of amount granted to use that maximizes the following utility:

$$U^L_{ijm} = \alpha^L + \delta^L_{jm}(X_{jm}, P_{jm}, \xi^L_{jm}, \beta^L) + V^L_{ijm}(Y_{ijm}, \eta^L) + \varepsilon^L_{im}. \tag{3.6}$$

Finally, conditional on borrowing, they will choose to default if the following utility is greater than zero:

$$U_{ijm}^F = \alpha^F + \delta_{jm}^F(X_{jm}, P_{jm}, \xi_{jm}^F, \beta^F) + V_{ijm}^F(Y_{ijm}, \eta^F) + \varepsilon_{im}^F. \tag{3.7}$$

Here X_{jm} are banks' observable attributes, P_{jm} are the posted interest rates mentioned above, ξ_{jm} are banks' unobservable (to the econometrician) attributes, and Y_{ijm} are firms' observable characteristics. We assume that ε_{ijm}^D is distributed as a type 1 extreme value, following the literature on demand estimation for differentiated products (Berry [1994], Berry et al. [1995]). We let the random coefficient of the demand's constant term $\alpha_i^D = \bar{\alpha}^D + \sigma_D \nu_i$, with $\nu_i \sim N(0,1)$, to be jointly normally distributed with ε_{im}^L, and ε_{im}^F, such that:

$$\begin{pmatrix} \alpha^D \\ \varepsilon^L \\ \varepsilon^F \end{pmatrix} \sim N\left(\begin{pmatrix} \bar{\alpha}^D \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & 0 & \rho_{DF} \\ 0 & \sigma_L^2 & \rho_{LF} \\ \rho_{DF} & \rho_{LF} & 1 \end{pmatrix} \right). \tag{3.8}$$

We interpret a positive correlation between the firm specific unobservables driving demand and default (ρ_{DF}) as evidence of adverse selection. The intuition is that if the unobservables that drive demand are positively correlated with the unobservables that drive default, then riskier firms are more likely to demand. The idea behind the identification of the correlation between α_i^D and ε_{im}^F is the following. If we observe a firm taking out a loan, while the model tells us that this firm should be unlikely to take the loan, then this is a "high α_i^D" firm. A positive correlation of α_i^D with ε_{im}^F is evidence of adverse selection.

We interpret a positive correlation between the unobservables driving loan size and default (ρ_{LF}) as evidence of moral hazard. The intuition is that if the unobservables that drive the choice of how much credit to use are positively correlated with the unobservables that drive default, then riskier firms will use more credit. We define this as moral hazard because the decision on how much loan to use is an action taken after the borrower and lender have agreed on the contract terms. With this definition of moral hazard we are trying to capture the case in which a risky firm (high ε_{im}^F), before signing the contract, already knows that due to its high ε_{im}^L it will use a higher share of the loan. However, our definition cannot rule out the case in which two ex-ante equally risky firms take the same loan, and one of them is hit by a negative shock after the contract has been signed. This shock increases ε_{im}^L for the

firm that was hit, forcing it to use more of the loan, but not due to moral hazard. [10] This identification strategy allows us to recover adverse selection and moral hazard parameters that are common across banks and markets, not bank or market specific.

On the supply side, we let banks set their interest rates competing à la Bertrand Nash. We assume that bank j's profits are given by the sum of profits made with each subset of borrowers' types k:

$$\begin{aligned} \Pi_{jkm} &= (P_{jkm} - MC_{jm})Q_{jkm}(1 - F_{jkm}) - MC_{jm}Q_{jkm}F_{jkm} \\ &= P_{jkm}Q_{jkm}(1 - F_{jkm}) - MC_{jm}Q_{jkm}, \end{aligned} \tag{3.9}$$

where Q_{jkm} and F_{jkm} are bank's expectation of demand and default. In particular, Q_{jkm} is given by the model's market shares and the expected loan size, and F_{jkm} is the average default rate for the borrowers of type k that bank j lends to in market m, following Assumption 1. P_{jkm} is the price of the loan $(1 + r_j)$. MC_{jm} are the bank's marginal costs, which we assume to be constant at the bank-market level. The first order conditions of this profit function deliver the following pricing equation:

$$\begin{aligned} & P_{jkm} = MC_{jm} - \frac{Q_{jkm}}{Q'_{jkm}} + AIC_{jkm}, \\ \textit{with} \quad & AIC_{jkm} = \frac{MC_{jm}F_{jkm} + MC_{jm}F'_{jkm}\frac{Q_{jkm}}{Q'_{jkm}} - F'_{jkm}\left(\frac{Q_{jkm}}{Q'_{jkm}}\right)^2}{1 - F_{jkm} - F'_{jkm}\frac{Q_{jkm}}{Q'_{jkm}}}. \end{aligned} \tag{3.10}$$

Note that the equilibrium price depends on marginal costs and markup $\frac{Q_{jkm}}{Q'_{jkm}}$, as in a standard Bertrand-Nash model with differentiated products, but also on a term defined as the Asymmetric Information Cost (AIC_{jkm}). This term is a function of default probability F_{jkm}, derivative of default with respect to prices F'_{jkm}, bank's markup and marginal costs. Both marginal revenues and the default probability determine the shape of the banks' profit function, driving it in different directions. The effect of an increase in interest rates increases on one hand the marginal revenues from borrowers that don't drop out, but on the other hand increases also costs from defaults, in the presence of adverse selection.

[10] We don't have a clear economic interpretation of the correlation between demand and loan size unobservables, so at the moment we are setting the correlation between them to zero for simplicity. We are planning to estimate it in future versions of the model.

3.4.2 Dynamic Game of Entry, Exit, and Investment

The players of the dynamic game are all the J banks entering/exiting, investing and competing in M markets. These banks decide their static pricing strategy and their dynamic entry/exit and investment decisions in every market in every year t. We assume that banks group borrowers into types, based on the observable characteristics they have access to. We also assume that banks set posted interest rates for each type of borrower because they cannot observe their individual probability of default, but can observe just the average default rate of each type, as in Stiglitz and Weiss [1981].

We focus on the years right after entry deregulation (late 1980s until 1998), during which the industry experienced a steady growth in entry rates, as well as a substantial increase in the number of branches. This means that the model environment is non stationary, as the number of players and the stock of branches in each market are increasing over time. To give a sense of the patterns in the data, Figure 3.5 shows the mean and standard deviation of the number of banks per province per year, whereas Figure 3.6 presents mean and the standard deviation of the number of branches per bank per province per year.[11] The market is experiencing a clear transition starting at the end of the 1980s, when branching deregulation took place.

[11] Note that these figures are based on the branch network dataset, which includes all banks, not only the 90 main banks in the CR data.

Figure 3.5: Average and Std Deviation of Number of Banks per Province per Year

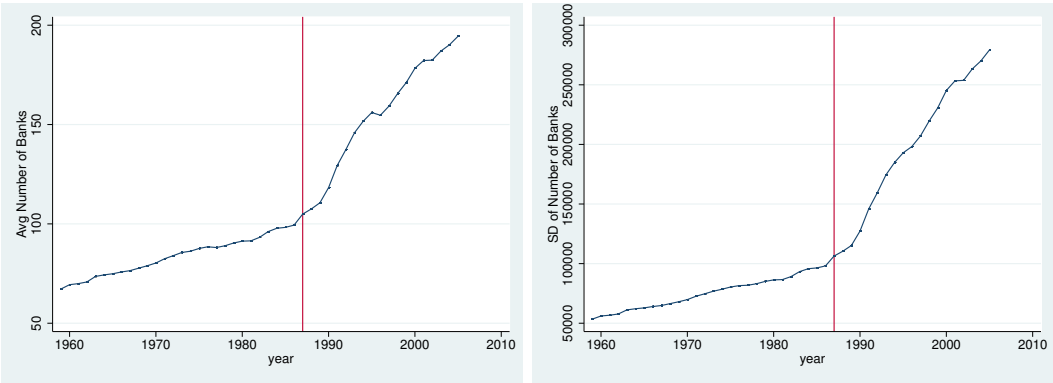
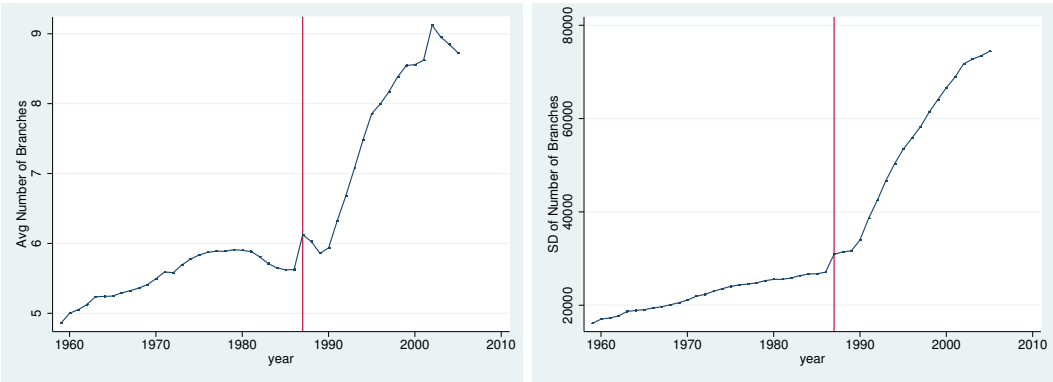


Figure 3.6: Average and Std Deviation of Number of Branches per Bank per Province per Year



The non stationary environment means that we cannot rely on the stationary Markovian structure that characterize the notion of Markov Perfect Equilibrium (MPE) in Ericson and Pakes [1995]. Therefore also two-step methods as Bajari et al. [2007] cannot be applied, as the range of the state variables that determine policy and value functions is changing over time. An alternative solution would be computing the equilibrium strategies and value functions at every period, with a clear computational problem due to the curse of dimensionality. We model instead this short run dynamic behavior of the industry using the notion of Nonstationary Oblivious Equilibrium (NOE), proposed by Weintraub et al. [2008a]. This equilibrium concept has been proven to approximate well MPE for industries with a large number of firms that are experiencing short run transitional dynamics after a shock or policy change, converging to a new stationary equilibrium in the long run. This approach is an extension to the Oblivious Equilibrium (OE) concept introduced by Weintraub et al. [2008b], and Weintraub et al. [2010]. Given the large number of players, NOE assumes that firms make decisions based only on their own state variables and on the deterministic average industry state, meaning that individual firms' actions don't affect directly the other firms' decisions.[12] This assumption well suits our case, as we have on average 26 banks per market.

We let the industry evolve over discrete time periods and an infinite horizon, in line with Ericson and Pakes [1995]. Time periods are indexed by $t \in \{0, 1, 2, ..., \infty\}$, and there is a discount factor $\beta \in (0, 1)$. The number of incumbent banks at time t in market m is defined as n_{tm}. A relevant bank-specific state variable is experience $e_{jtm} \in \{0, 1, 2, ..., E\}$, which varies over markets and time, and influences borrowers' demand and default rate. The industry state s_{tm} is a vector that specifies the number of banks at experience level e in period t in market m. We also define the state of competitors of bank j as $s_{-jtm}(e) = s_{tm}(e) - 1$ if $e_{jtm} = e$, and $s_{-jtm}(e) = s_{tm}(e)$ otherwise. Similarly, n_{-jtm} is the number of competitors of bank j.

Incumbent banks can decide to invest A_{jmt} through branching to improve their experience level, but experience can also depreciate due to organizational forgetting. We assume that branching decisions are market specific, even though banks operate across different markets. This assumptions derives from the empirical evidence of local banking competition (Degryse and Ongena [2005]), which implies that it is reasonable to think that local managers play an important role in taking deci-

[12] Benkard et al. [2013] recently developed an extension to this framework called Partially Oblivious Equilibrium. This allows for some dominant firms playing Markovian strategies, whose states are always monitored by all other firms in the market, and for a competitive fringe playing Oblivious strategies but monitoring the dominant firms.

sions about local investment, as they have a better knowledge of the market.[13]. We construct the experience state variable as a function of the number of branches that a bank has in a market and the number of years that those branches have been in that market. We let experience influence borrowers' demand, as firms might have a preference for larger and well established networks of branches. We also allow experience to have an impact on borrowers' default rate, as experienced banks attract a better pool of firms. We already provided empirical supporting evidence in the reduced form section for this assumption.

Following this literature on learning by doing however doesn't allow us to model in detail the way banks reduce, if any, the extent of asymmetric information and learn about their borrowers. It does allow us to find a reduction in banks' borrowers' default rates as the bank accumulates experience in a market overtime. The reduction in asymmetric information is one of the possible explanation of this reduction in default rates, but we cannot rule out other possible causes, like increased bank efficiency or more better matching in firm-bank relationships. One of the possible ways to identify a reduction in asymmetric information would be to model the correlation coefficients as a time-varying process, such as an autoregressive one, dependent on the year of the firm-bank relationship. This would allow us to test whether asymmetric information reduces as firm-bank relationships evolve, quantifying how much it contributes to the decline in default rates.[14]

Following Weintraub et al. [2010] and Pakes and McGuire [1994], we define the transition probability of experience as:

$$P(e'|e,A) = \begin{cases} \frac{(1-\delta)\theta(A+e)}{1+\theta(A+e)} & \text{if } e' = e+1 \\ \frac{1-\delta+\delta\theta(A+e)}{1+\theta(A+e)} & \text{if } e' = e \\ \frac{\delta}{1+\theta(A+e)} & \text{if } e' = e-1 \end{cases} \tag{3.11}$$

where θ is the branching efficiency parameter, and $\delta \in [0,1]$ is the probability that banks' experience depreciates, capturing organisational forgetting. Note that under this specification of the transition probability banks can increase their stock of experience even without opening new branches. This is because experience grows also when new firms borrow from existing branches. We model banks' investment cost as $C(A_{jtm}, \psi_{jtm}) = \max\{(1+\gamma\psi_{jtm})A_{jtm,0}\}$, in line with Qi [2013], where $\psi_{jtm} \in N(0,1)$ is an IID investment shock, private information of each bank, and γ is an investment cost.

[13] Aguirregabiria and Ho [2012] make a similar assumption about decisions of local managers in the airline industry, in the context of a dynamic entry and exit game.

[14] Extending the model in this direction is scope for future research.

Each period incumbent banks can decide to exit a market if a privately observed sell-off value ϕ_{jmt} is greater than the continuation value. We express this decision as $\chi_{jmt} = 1$. Similarly, every period new firms can enter the market paying a sunk cost κ, and they all enter with zero experience. This action is denoted by $\epsilon_{jmt} = 1$.

The timing of the game is the following:

- All banks observe the industry state s_{tm}
- Incumbent banks receive a private draw from the distribution of scrap values of exit and decide to leave the market or not.
- Potential entrants receive a draw from the distribution of entry costs and make their entry decision.
- Incumbent banks receive a draw from the distribution of investment cost and make their branching decision.
- Incumbent banks compete over interest rates.
- Incumbent banks accumulate experience and exit if they decided so. Potential entrants enter the market.
- The industry evolves to s_{t+1m}.

3.4.3 Equilibrium

We approximate a Markov Perfect Equilibrium using the concept of Nonstationary Oblivious Equilibrium, which accommodates a short run dynamic industry behavior assuming that there will be a stationary oblivious equilibrium in the long run, once the transition period is over. We follow the set up outlined by Qi [2013]. This notion adapts well to our case, given the large number of banks in each market and then nonstationary environment that characterises the period we're looking at. One of the main assumptions in a NOE is that banks' strategies in terms of entry (ϵ_t), exit (χ_t), and investment (ι_t) do not depend on the industry state s_{tm}. This is because the trajectory of the expected industry state $\tilde{s}_{tm}$, given an initial industry state s_{0m}, evolves deterministically. Hence, the policy functions will only depend on the relevant state variables e_{jmt} and ψ_{jmt}. The incumbent's strategy is defined as $\mu_t = \{\chi_t(e_{jmt}), \iota_t(e_{jmt}, \psi_{jmt})\}$. From now on we suppress the bank-market subscripts for simplicity.

Given all banks follow the same strategy profiles μ, ϵ and an initial industry state s_0, we denote as $\tilde{s}_{\{\mu,\epsilon,s_0\},t}$ the expected industry state at time t, and define banks' experience transition probability function as:

$$\mathbb{P}(e'|e,\psi,\mu_t) = P(e'|e,\iota_t(e,\psi))(1-\exp(-\chi_t(e)/K), \tag{3.12}$$

where the second term on the right hand side is the probability of a bank not exiting, under the assumption of the scrap value ϕ being drawn from an exponential distribution with mean K. We can define the expected experience transition probabilities as $\mathbb{EP}_{\mu_t}(e,e') = \mathrm{E}_{\psi}[\mathbb{P}(e'|e,\psi,\mu_t)]$, so that the sequence of industry states will be:

$$\tilde{s}_{t+1}(e) = \begin{cases} \sum_{k=0}^{E} \mathbb{EP}_{\mu_t}(k,e)\tilde{s}_t(k) + \epsilon_t & \text{if } e = e_{ent} \\ \sum_{k=0}^{E} \mathbb{EP}_{\mu_t}(k,e)\tilde{s}_t(k) & \text{otherwise} \end{cases} \tag{3.13}$$

where e_{ent} is the level of experience of a new entrant. We define a nonstationary oblivious value function for period t for an incumbent bank at experience level e, conditional on its own strategy μ', its competitors' strategies μ and entry rates ϵ, as:

$$\tilde{V}_t(e|\mu',\mu,\epsilon,s) = \Pi_t(e,\tilde{s}_t) + E_{\phi}\big[\max\{\phi_t, E_{\psi}[\widetilde{VC}_t(e,\psi|\mu',\mu,\epsilon)]\}\big], \tag{3.14}$$

where $\widetilde{VC}()$ is the incumbent's continuation value conditional on not exiting, expressed as:

$$\widetilde{VC}_t(e,\psi|\mu',\mu,\epsilon) = \max_{\iota} -C(\iota,\psi) + \beta E_{e'}[\tilde{V}_{t+1}(e'|\mu',\mu,\epsilon)]. \tag{3.15}$$

A NOE consists of a strategy profile $\tilde{\mu}$ that maximises the incumbent's value function and of a strategy profile $\tilde{\epsilon}$ that maximises the entrant's value function, such that $\epsilon_t = 1$ if:

$$V_{ent}(e_{ent}) = \beta\tilde{V}_{t+1}(e_{ent}|\tilde{\mu},\tilde{\mu},\tilde{\epsilon}) \geq \kappa \tag{3.16}$$

3.5 Estimation

We estimate the static part of the model as in Crawford et al. [2013], which explains the details of the econometric model. We use the demand, loan size and default probabilities to construct the simulated maximum likelihood that allows us to recover the parameters in $\eta = \{\eta^D, \eta^L, \eta^F\}$,[15] and the correlation coefficients ρ_{DF} and ρ_{LF}:

$$\log L = \sum_i \log(\Pr^D_{ijm}) D_{ijm} + \sum_{i \in D} \Big[\log(\Pr^L_{ijm}) + \log(\Pr^F_{ijm}) F_{ijm} + \log(1 - \Pr^F_{ijm})(1 - F_{ijm}) \Big], \tag{3.17}$$

where $\Pr^D_{ijm}$ is probability that borrower i in market m chooses lender j, $\Pr^L_{ijm}$ is the probability of observing a utilization of L_{ijm}, and $\Pr^F_{ijm}$ is the probability of default conditional on taking a loan. D_{ijm} is the actual bank choice from the data, and F_{ijm} is the default from the data. In order to estimate the parameters $\beta = \{\bar{\alpha}^D, \alpha^L, \alpha^F, \beta^D, \beta^L, \beta^F\}$ we need an additional step. Given some instruments Z_{jm}, we recover β using instrumental variables.

To estimate the dynamic parameters $\Theta = \{\theta, \delta, \gamma, \phi, \kappa\}$, we follow a simulated method of moments similar to other applications of oblivious equilibrium, as Xu [2008] and Qi [2013]. These parameters include branching efficiency θ, experience depreciation δ, cost of investment γ, exit scrap value ϕ, and sunk cost of entry κ.

[15] In this version of the model we are not estimating the standard deviation of the random coefficient of the constant term in demand, setting it to $\sigma_D = 1$. This is due to the well known identification problem of these coefficients in Berry et al. [1995], explained in Berry et al. [2004] and Train and Winston [2007]. We are working on incorporating second preferred choices into the model to guarantee better identification and be able to estimate that parameter.

3.6 Counterfactuals

We compare the estimated model against two counterfactual scenarios. One is the case of no borrower-lender asymmetry of information, the other is the case of reduced information gap between experienced and unexperienced banks. In the first case, we set the correlation coefficients that identify asymmetric information to zero. In the second case, we allow new entrants to have a higher initial experience stock. Doing this, we want to explore the consequences in terms of welfare, credit rationing, and market structure of these two dimensions of information asymmetry. The type of policy that could implement these reduction could be a regulation that allows for more information sharing between banks, or also a policy that restricts the number of incumbents that a market could have. One of the questions we try to answer is whether concentration can improve welfare in the presence of asymmetric information.

3.7 Conclusion

This paper investigates the effect of asymmetric information on market structure in the Italian banking industry. Using a detailed and highly representative dataset of contracts between borrowers and lenders over 11 years, we show that incumbent banks accumulate an informational advantage over potential entrants, and this advantage affects banks' entry and exit decisions. Hence, we develop a dynamic structural model of banks' entry, exit and investment through branching that allows banks to accumulate experience and improve the quality of their borrowers over time, and gives incumbent banks an informational advantage over new entrants. We incorporate into this framework a companion static model of firms' demand for credit, loan size, default, and banks' pricing from Crawford et al. [2013], which allows for borrower-lender asymmetric information. We assume that banks follow nonstationary oblivious strategies, as defined by Weintraub et al. [2008b] and Benkard et al. [2013], given the large number of players in each market and the nonstationary environment of the post-deregulation period.

We provide reduced form evidence of learning by branching, showing that banks improve the quality of their borrowers as they accumulate experience over the years. We also show that entry is negatively correlated with concentration of incumbents' experience, and that exit is negatively correlated with a bank's own experience level. We propose two counterfactual policy experiments to quantify the effects of information asymmetries between borrowers and lenders and between incumbents and potential entrants on market structure, welfare, and credit rationing. In the first case, we do so computing a new equilibrium once we eliminate what we define as adverse selection and moral hazard. In the second case, we allow newly entering banks to have a higher level of experience with respect to the actual data, and compare the new equilibrium outcomes that emerge.

Chapter 4

Market Structure and Multi Market Contact

Market Structure and Multi Market Contact*

Lorenzo Ciari[†], Nicola Pavanini[‡]

June 2014

*We thank Gregory Crawford, Carlos Noton, Andrea Pozzi, Giancarlo Spagnolo, and Mike Waterson for helpful comments.

[†]EBRD, Ciaril@ebrd.com

[‡]University of Warwick, N.Pavanini@warwick.ac.uk

Abstract

This paper measures the impact of endogenous multi market contact on entry decisions of Italian national banks. We develop a static model of market structure with incomplete information as in Seim [2006], allowing for global players' heterogeneity and spatial correlation of entry decisions across different local markets. The framework is estimated based upon the expansion of the three main Italian commercial banks, Unicredit, Intesa San Paolo and Monte dei Paschi, which have been progressively gaining control of the market after 1990's entry deregulation. Preliminary results show that multi market contact enhances banks' profitability, suggesting that it might facilitate implicit collusion as in Bernheim and Whinston [1990].

4.1 Introduction

The effects of multi market contact on a firm's strategic decisions have been widely investigated, both theoretically and empirically, to detect a possible channel for tacit collusion. Most of the literature has focused on the implications for competitiveness, taking market structure as given. However, besides the collusive effects on whichever dimension firms are competing upon, multi market contact can also have a significant impact on firms' entry decisions.

The first paper documenting a potentially anticompetitive outcome in the presence of linked oligopoly was Edwards [1955], which introduced the idea that competition between large rivals whose interests touch at many points may "convert a warfare into total war", as the author says, supporting the incentive to a live and let live strategy. Feinberg [1984] contributed to the theoretical foundations of multi market contact, analyzing the effects of cross-market mutual forbearance as an extension of standard oligopoly theory. Later, Bernheim and Whinston [1990] formalized the results of collusive gains from multi market contact only for the cases of either non identical firms, or non identical markets, or differentiated products. Spagnolo [1999] found that multiple links always facilitates collusion when the firms' objective functions are strictly concave and market supergames are "interdependent".

Most of the empirical literature finds a positive and significant relationship between multi market contact and profitability. Hughes and Oughton [1993] look at the effect of diversification and multiple links on profitability across 134 UK manufacturing industries, finding a clear positive effect. Jans and Rosenbaum [1993] investigate the impact of multi market contact on pricing in the US cement industry, showing a positive relationship between price-cost margins and number of contacts. Evans and Kessides [1994] concentrate on the US airline industry, finding that fares are higher in city-pair markets served by companies with extensive interroute contacts. The authors claim that these results are driven by airlines avoiding aggressive pricing for fear of retaliation in other city-pairs. On this basis, Ciliberto and Williams [2013] propose a structural approach to show that multiple contacts between US airlines facilitate collusion. They construct a model of oligopolistic behavior, where some conduct parameters are functions of pair-specific multi market contacts. The main result is that carriers with numerous contacts can sustain near-perfect cooperation in setting fares. Also a recent paper by Chicu and Ziebarth [2013], applied to the US cement industry, finds that multi market contact fosters tacit collusion and higher prices. The authors present a novel measure of contact that accounts for capacity utilization.

Relaxing the underlying assumption of exogenous market structure leaves scope for several other relevant questions. How does a multi market firm's entry decision depend on the number of contacts it has with the potential rivals entering the same market? Do firms prefer to have many links with their rivals or do they avoid each other? In one case they might find it profitable to establish many contacts, as this would facilitate collusion and make any form of retaliation more costly. In the other case, they might prefer dividing up the markets in spheres of influence, or might value the option of pricing aggressively in some specific markets without experiencing widespread price wars. In a recent paper Byford and Gans [2014] extend the theory model of Bernheim and Whinston [1990] allowing for entry and exit with multimarket contact. The authors show that there can be an equilibrium with collusion at the extensive margin, where firms collude by avoiding entry into each others' markets. This implies that a system of mutual forbearance might be optimal for firms, in a strategy to divide and conquer separate geographical markets. We propose to address these questions developing a static model of market structure with incomplete information as in Seim [2006], introducing endogenous multi market contact creation. We model the entry decisions of heterogeneous global players, allowing for spatial correlation of entry across different markets through a multi market contact index. We experiment with various indices, but mostly concentrate on contacts in neighboring markets rather than all possible existing contacts. This is motivated by the important role of distance in banking and by the local nature of banking competition, as described in Degryse and Ongena [2005]. In this paper we focus on the relationship between profitability and number of contacts, controlling for demand and cost effects using market specific characteristics.

The case of banks' expansion in Italy is a good candidate for this analysis for several reasons. After entry deregulation at the beginning of the 90s, there has been a steady consolidation process. By 2005 there had been a 152% growth in the national Herfindahl-Hirschman Index (HHI) in terms of branches, and the 3 main players Unicredit, Intesa San Paolo and Monte dei Paschi's combined shares passed from 8% in 1990 to around 30% in 2005. Given the important role played by these 3 banks, we focus our analysis on their expansion patterns, considering the remaining local and savings banks as a competitive fringe. The literature on multi market contact in the Italian banking sector presents conflicting evidence. For the 1990-1996 period, De Bonis and Ferrando [2000] test how multi market links with other banks across provinces affect the change in bank-province-year specific loan market shares, finding a positive effect. They also find that multiple contacts are negatively correlated with lending rates. These two results contradict the multi

market contact hypothesis. On the other hand, Coccorese and Pellecchia [2009] show that in later years (2002-2005) market-level and firm-level profitability, measured in terms of return on assets, is positively correlated with the average number of contacts among banks. Finally, Molnar et al. [2011] propose a structural model of demand and supply of deposit retail services, testing banks' conduct under Bertrand-Nash, partial or perfect collusion based on banks' contacts. They find evidence of partial collusion, especially for banks with more contacts.

We present reduced form and structural evidence that shows how multi market contact has a positive influence on entry probability, with different effects for different banks. We find a stronger impact for closer neighboring markets and for lower share of branches of local players. We also allow for homogeneous or heterogeneous effects of multi market contact across the three different banks. In one specification, we impose the same coefficient for the three of them. In a second one, we allow each bank to have a different effect of multi market contact on their profitability. In the last specification, we allow for bank-to-bank specific multi market contact coefficients. We find a substantially positive effect of multi market contact in the first two cases.

The contribution of the paper is twofold. First, we expand a static entry model with incomplete information to allow for endogenous multi market contact creation. We investigate a novel aspect for the literature on static entry in general (Bresnahan and Reiss [1991], Berry [1992], Mazzeo [2002], Toivanen and Waterson [2005]), as well as for its banking applications (Cohen and Mazzeo [2007]). Second, we contribute to the empirical banking literature, where there is little evidence of the effect of multi market contact on entry (Fuentelsaz and Gomez [2006]). Differently from most of the papers on linked oligopoly in Italy, we define the geographical markets in term of local labor systems (LLS) instead of provinces. These local markets are comparable in size to the Labor Market Areas in the US, used for most of the entry literature in the US, and are defined according to integrated economic areas and commuting patterns between councils. For these reasons, the LLS are likely to represent more accurately the area where banks compete.

This paper is organized as follows. Section 2 describes the dataset and the local markets, section 3 designs the framework and section 4 outlines the econometric specification. The reduced form and structural results are described in section 5, section 6 concludes.

4.2 Data and Local Markets

We apply our model to the case of the main Italian banks' expansion up to 2005. The dataset employed consists of the yearly branching network for the population of Italian banks between 1959 and 2005 at the city council level. It also includes all the merger and acquisition patterns, mostly concentrated at the end of the 1990s. We will however only make use of the market structure in 2005 for our application.[1]

The Italian banking sector has been heavily regulated since the 1936, in response to the effects of the 1929 great depression on the credit market. The solution adopted by the Italian government was to strictly discipline the sector both in terms of branching and in terms of types of credit instruments that banks were allowed to provide, with a specific aim for geographic and scope risk diversification (Polsi [2000]). Different banks were in fact restricted to different functions in terms of length of credit relationships, economic sectors, as well as territorial areas of expansion.

The European economic integration process fostered a substantial deregulation of the Italian credit market, that took place between 1990 and 1993. The reforms implemented were aimed at favoring the development of universal banks that could compete against each other, removing all the scope and territorial expansion restrictions that had been in place for over 50 years. In the years that followed the sector experienced a consolidation process that led the three main credit institutions, Unicredit (UC), Intesa San Paolo (SP), and Monte dei Paschi di Siena (MPS), to control around 30% of the branches in the country by 2005. For this reason, we focus our attention on these three national banks, and regard local, cooperative, and savings banks as a competitive fringe.

Differently from most of the entry literature on Italian banking, we don't consider the province as a local market, but rather the Local Labour Systems (LLS). These are defined by the Italian national statistical institute (ISTAT) as urban agglomeration units of city councils geographically and statistically comparable. They are used by ISTAT to investigate the socio-economic structure of Italy from a local perspective, and are constructed based on commuting patterns between councils, just like the LMA in the United States. This definition of local markets, based on economic contiguity among bordering local areas, appears to be more appropriate than that of provinces, which represent bigger administrative units. Following Cohen and Mazzeo [2007], the required characteristics for a local banking market are: (1) consumers in the defined geographic markets should not typically use depository

[1] We will explain this choice in detail in the results' section.

institutions outside of their area and (2) distinct or overlapping submarkets should not exist within the defined geographic markets. Table 4.1 presents some descriptive statistics on size, population, and economic activity of the LLS. Table 4.2 describes the LLS in terms of branch networks, both for national and local banks.

The 1990 entry deregulation gave rise to a large wave of entry and mergers and acquisitions that increased substantially the levels of concentration in local markets. During this period, the three main national banks gained a relevant market share and expanded their network of reciprocal contacts across LLSs. Figure 4.1 shows the evolution of the Herfindahl-Hirschman Index (HHI) in terms of branches over the 1990-2005 period, measured every five years at the national, province and local labor system level.

The data suggests a clear diverging pattern whether we look at the national or local dimension. At the national level, the concentration of the industry significantly increased, with the HHI going from 0.0116 in 1990 to 0.0292 in 2005 (a growth rate of 152%). As a result of the mergers and overall consolidation, few banks acquired control of the market at the country level. However, when we turn to the local dimension concentration decreased constantly over the period considered, both for provinces and LLSs. So, if on one side many small players exited the market and few big players gained control, local markets become apparently more contestable. Figure 4.2 reports the joint market share of the three big banks at the national and LLS level. The combined shares at the country level grow from 8.77% in 1990 to 28.31% in 2005, and similarly at the local level. Figure 4.3 shows how the number of contacts that the three main players have in local Italian markets increased from 1990 to 2005. In this figure, we measure multi market contacts as the sum of markets in which pairs of banks or the three main banks are both present with at least one branch.

Table 4.1: Summary Statistics on Local Labour Systems

Variable	N	Mean	Stand. Dev.	5th pctile	Median	95th pctile
N. of City Councils	686	11.81	13.35	2	7	36
Surface	686	439.25	349.44	88.64	353.80	1,152.17
Population	686	83,084.17	222,418.03	7,020	33,966.50	262,233
Employment Rate	686	43.18	7.20	32.70	42.95	54.10
Unemployment Rate	686	8.89	5.37	2.80	7.65	18.40
Manufacturing Units	686	861.19	2,125.92	55	283	2,923
Value Added	686	16,309.77	6,408.10	7,196.54	16,115.76	26,965.87
No Specialization	686	0.32	0.47	0	0	1
Urban	686	0.07	0.25	0	0	1
Non Manufacturing	686	0.19	0.39	0	0	1
Textile and Clothing	686	0.15	0.35	0	0	1
Other Made in Italy	686	0.19	0.39	0	0	1
Heavy Manufacturing	686	0.08	0.27	0	0	1

Note: One observation is a Local Labor System. Surface is measured in squared kilometers. Value added is measured in euros. The variables no specialization, urban, non manufacturing, textile and clothing, other made in Italy, and heavy manufacturing are binary.

Table 4.2: Summary Statistics on Local Labour Systems

Variable	Mean	Stand. Dev.	5th pctile	Median	95th pctile
Total Branches	46.19	125.14	3	17	149
1959-2005 % Branch Growth	340.15	291.10	50	259.41	900
Number of Main Banks	1.91	1.06	0	2	3
Number of Local Banks	8.36	11.61	1	5	24
Share of Main Banks' Branches	31.46	20.47	0	32.34	66.67
% Markets with Unicredit	0.71	0.45	0	1	1
Share of Unicredit's Branches	4.64	7.20	0	1.90	20
% Markets with San Paolo	0.66	0.47	0	1	1
Share of San Paolo's Branches	4.54	8.66	0	1.54	16.67
% Markets with MPS	0.53	0.50	0	1	1
Share of MPS's Branches	2.68	4.89	0	0.43	12.50

Note: One observation is a Local Labor System. Main banks are Unicredit, San Paolo, and MPS.

Figure 4.1: HHI level over the period 1990-2005

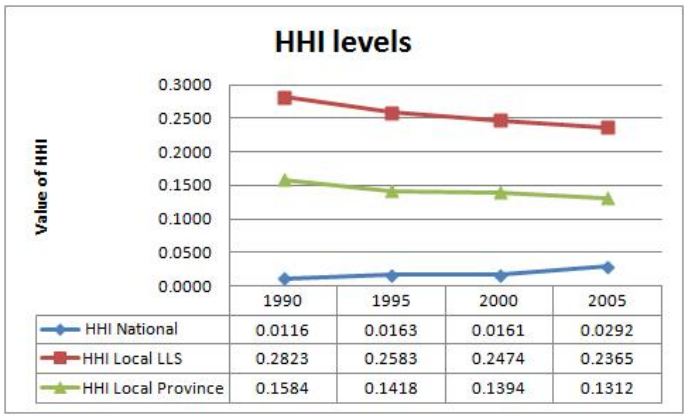
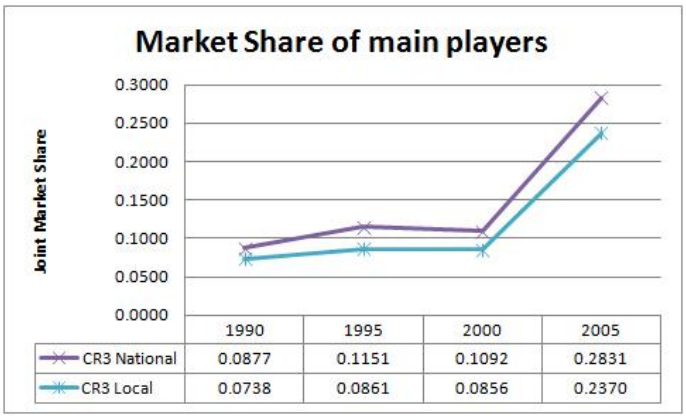
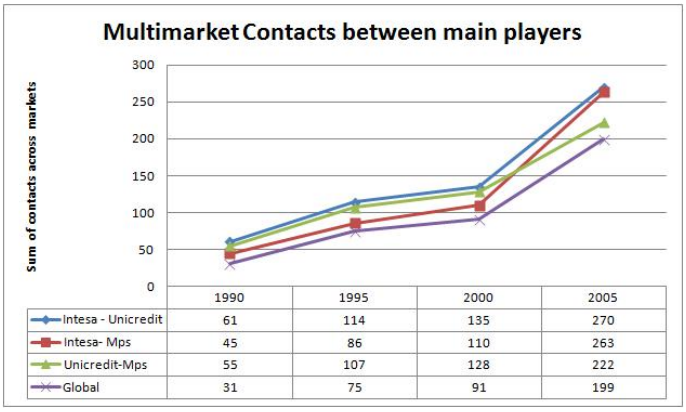


Figure 4.2: Joint market shares of three main players over the period 1990-2005



Note: Local concentration is at the Local Labor System level.

Figure 4.3: Multimarket contacts for the three main players over the period 1990-2005



Note: Multi market contact is the sum of Local Labor Systems where both banks are present.

4.3 The Model

The framework we construct builds on the work of Seim [2006], developing an entry model with endogenous multi-market contact creation. The main contributions of Seim [2006] were the introduction of incomplete information in static entry models, as well as local players' endogenous choice of spatial differentiation within a market. Assuming that players have private information about their own profitability simplifies the computation of equilibrium strategies, and provides a convenient solution to the problem of multiple equilibria, as shown with an existence and uniqueness proof that the author provides for a simple version of the game. Most importantly, it is reasonable to think that agents have a better knowledge of their own profitability compared to their rivals. This is particularly true for our application to the credit sector, where the private information could be about creditworthiness of a bank's own borrowers, as well as about the stock of market-specific soft information that the bank collects through lending relationships and years of experience in a local market.

If on one hand we maintain the assumption of incomplete information, we depart from Seim [2006]'s model in other dimensions. First, we concentrate on 3 main global players, instead of various local homogenous firms. This is a necessary condition to model multi market contacts, as we need players' identities to construct their links across markets. It also allows us to analyze heterogeneous effects of multi market contact on banks' profitability. Second, we don't let players choose their location within a market,[2] but instead allow their entry choices to be correlated across different markets through the multi market contact index. The idea is to capture player's rivalry effects through their expectations of reciprocal contacts across markets, where these contacts are defined over different radiuses of distance between a local market and its neighboring markets.

Let $d_{fm} = \{0, 1\}$ be the decision of a potential entrant firm $f = 1, ..., F$ to enter market $m = 1, ..., M$. We consider all the global players as potential entrants in each market. The payoff function of firm f will be:

$$\Pi_{fm} = \alpha + \xi_m + \beta X_m + \gamma Z_{fm} + \delta_f \sum_{g \neq f} \mathbf{1}[\Pi_{gm} > 0] * MMC_{fgm} + \varepsilon_{fm}, \qquad (4.1)$$

where ξ_m are unobservable (to the econometrician) exogenous market char-

[2] We are working on extending the model in this direction, as we have data on banks' location decisions within local markets.

acteristics, X_m is a vector of observable demand and cost characteristics of market m, and Z_{fm} is a vector of observable characteristics of firm f in market m. Note that we're allowing the coefficient δ_f of the multi market contact index MMC_{fgm} to vary by firm, assuming that MMC_{fgm} might have heterogeneous effects on firms' profitability. We can think of different degrees of heterogeneity across firms for the effect of multi market contact on profitability. The coefficient on the MMC index could also be homogeneous across banks (δ), or heterogeneous at the firm-pair level (δ_{gf}). We experiment with these three possible specifications in the estimation section. We are also conditioning the effect for firm f of multiple contacts with rival g on the actual presence of firm g in the market, through the indicator function $\mathbf{1}[\Pi_{gm} > 0]$.

The idiosyncratic component ε_{fm} of firm f's profits from operating in market m is private information of firm f, and independently and identically distributed over markets and firms. Assume ε_{fm} are IID draws from a type 1 extreme value distribution. This distributional assumption guarantees computational tractability, but requires firms to have a private information component in their profit function that is specific to each market, and not correlated across markets and firms. Given the local nature of retail banking (Degryse and Ongena [2005]), the IID condition can be viewed as the bank-market specific soft information, which determines a bank's own profits and is not disclosed to its rivals to preserve an informational advantage in a competitive market.

We do however allow for profit correlation across markets and firms through the multi market contact index MMC_{fgm}, which is defined as the ratio of the sum of all contacts in other markets that firm f has with each other potential entrant g in market m, over the total number of markets M minus one. This formula is quite standard in the empirical multi market contact literature, as in Evans and Kessides [1994] and Ciliberto and Williams [2013]. This differs from Seim [2006]'s work only in one dimension, because Seim allows rival's profitability (i.e. rivals' entry probability) to affect a firm's profitability, modeling a strategic interaction among firms within the same market, but still keeping the assumption of the shocks being uncorrelated across firms within a market. We also allow for that, as the multimarket contact index for firm f is weighted by the entry probability of each rival firm g, but on top we model the strategic interaction of the same firm f across different markets through the same index. A clear limitation of this index is that it is a reduced form representation of a rivalry effect, which includes at the same time these two dimensions of correlated profitability across markets and firms. This follows closely the empirical reduced form literature on multimarket contact, starting from Evans

and Kessides [1994], and would require a more coherent and sophisticated structural model to be able to separately identify the multimarket contact effect from other forms of strategic interactions among branches located in different markets of the same firm.[3]

We construct the multimarket contact index as:

$$MMC_{fgm} = \frac{1}{M-1} \sum_{k \neq m} 1[g \textit{ and } f \textit{ both active in market } k]. \tag{4.2}$$

Each firm forms an expectation of each rival's optimal location choices. Therefore, the expected profit of entering in market m is:

$$\begin{aligned} E[\Pi_{fm}] &= \alpha + \xi_m + \beta X_m + \gamma Z_{fm} + \delta_f \sum_{g \neq f} \mathbf{1}[E(\Pi_{gm}) > 0] * E[MMC_{fgm}] + \varepsilon_{fm} \\ &= E[\overline{\Pi}_{fm}] + \varepsilon_{fm}. \end{aligned} \tag{4.3}$$

We can define the probability that competitor $g \neq f$ chooses market m as:

$$p_{gm}(d_{gm} = 1 \mid \xi, X, \theta) = Pr(E[\overline{\Pi}_{gm}] + \varepsilon_{gm} > 0), \tag{4.4}$$

for every m, where $\theta = (\alpha, \beta, \gamma, \delta)$. All firms will have an expectation on rivals' entry probabilities:

$$E(\Pi_{gm}) > 0 = p_{gm}, \tag{4.5}$$

[3] A more structural approach is scope for future research.

whereas the expected level of multi-market contact of firm f with firm g will be:

$$E[MMC_{fgm}] = \frac{1}{M-1} \sum_{k \neq m} p_{gk} * p_{fk}. \tag{4.6}$$

Hence, we can construct a system of $F * M$ equations that defines the equilibrium location conjectures as a fixed point of the mapping from the firm's conjecture of its rivals' strategies into its rivals' conjectures of the firm's own strategy as follows:

$$p_{fm} = \frac{\exp(\alpha + \xi_m + \beta X_m + \gamma Z_{fm} + \delta_f \sum_{g \neq f} p_{gm} \frac{\sum_{k \neq m} p_{gk} * p_{fk}}{M-1})}{1 + \exp(\alpha + \xi_m + \beta X_m + \gamma Z_{fm} + \delta_f \sum_{g \neq f} p_{gm} \frac{\sum_{k \neq m} p_{gk} * p_{fk}}{M-1})} = \frac{\exp(\xi_m) \exp(\widetilde{\Pi}_{fm})}{1 + \exp(\xi_m) \exp(\widetilde{\Pi}_{fm})}. \tag{4.7}$$

Existence and uniqueness of this Bayesian Nash equilibrium is discussed in detail in Seim [2006].

4.4 Econometric Specification

In order to estimate the model following Seim [2006], we assume that the expected number of entrants predicted by the model equals the actual entrants in the data Ω_m. This is done to simplify the estimation of the highly nonlinear entry model in 4.7, adjusting the market specific unobservable ξ_m until the expected number of entrants equates the actual number of entrants in each market:

$$\sum_f p_{fm} = \Omega_m, \tag{4.8}$$

where Ω_m is the total number of firms that have entered in market m, according to the data. The method of using a market specific unobservable to induce equivalence between predicted and actual number of entrants follows from Berry [1994] and Berry et al. [1995], who applied it to equate predicted and actual market shares in a differentiated products demand model. We can determine the market specific unobservable ξ_m solving the system of equations 4.7 and 4.8, which in our case, where the maximum number of entrants considered is $F = 3$, can be expressed as:

$$\frac{\exp(\xi^m)[\exp(\widetilde{\Pi}_1^m)]}{1+\exp(\xi^m)[\exp(\widetilde{\Pi}_1^m)]} + \frac{\exp(\xi^m)[\exp(\widetilde{\Pi}_2^m)]}{1+\exp(\xi^m)[\exp(\widetilde{\Pi}_2^m)]} + \frac{\exp(\xi^m)[\exp(\widetilde{\Pi}_3^m)]}{1+\exp(\xi^m)[\exp(\widetilde{\Pi}_3^m)]} = \Omega_m. \tag{4.9}$$

In order to solve for $\exp(\xi)$, this can be further simplified to the following market specific F^{th} order polynomial:

$$\sum_{f=1}^{F} (f - \Omega_m)[h_m(f)] \exp(\xi)^f = \Omega_m, \tag{4.10}$$

where $h_m(f)$ is a function that varies depending on f.[4] This polynomial can have at most F solutions. We select the solution that maximizes the likelihood function compared to the other solutions.

The likelihood function is based on the probability of entry, conditional on the distribution of the market specific unobservable component, and on the probability of observing the ξ_m realization that equates predicted and actual number of entrants

[4] See the appendix for the derivation of this polynomial, i.e. all the steps to go from 4.9 to 4.10, and what exactly is the function $h_m(f)$.

in market m:

$$\mathrm{L}(\theta) = \prod_m p(d_{fm}|\xi_m, X_m, \Omega_m) g(\xi_m|X_m, \Omega_m). \tag{4.11}$$

Under the assumption that $\xi_m \sim N(\mu, \sigma^2)$, the normal density $g(\xi_m)$ of each observation ξ_m will allow us to estimate the mean and standard deviation of the distribution of market specific unobservables.

4.5 Results

The empirical application we propose looks at the effect of multi market contact on the 2005 market structure of the Italian banking sector. The idea of using the single most recent cross section of data comes from the example of Bresnahan and Reiss [1991]. Since their seminal contribution, several papers on entry have been actually modeling a two-stage game of market structure, where firms simultaneously decide to enter in the first period, and then compete on prices or quantities (Berry [1992], Mazzeo [2002], Toivanen and Waterson [2005], Cohen and Mazzeo [2007]). Most of the times the second stage is not modeled structurally, and its outcome is approximated using a reduced form profit function. A clear limitation of these models is the lack of dynamics, that requires to assume that the industry is in a long run stable equilibrium. We follow this approach, assuming that after 15 years of entry deregulation the industry has reached a stable and consolidated market structure. [5]

We focus primarily on the three main national banks, Unicredit, San Paolo, and Monte dei Paschi di Siena, treating local, cooperative, and saving banks as a competitive fringe. Following the model specification presented above, we construct a reduced form profit function based on observable market and bank characteristics, as well as on a multi market contact index. The market observable attributes considered are surface, population, value added and employment rate to control for market-specific exogenous demand or cost shifters. We also include the bank-market specific distance from the bank's national headquarter to proxy for any possible benefits (costs) of opening a branch close to (far from) the main decisional centre of each credit institution. One limitation of this model is that it consider entry by merger or acquisition as a market-specific decision, even though most of these consolidations involve acquiring branches in several markets. Perez-Saiz [2013] describes the problems in identifying the fundamental primitives of an entry model when the options of greenfield entry and entry by acquisition are not considered as different strategies. We are aware of these possible drawbacks, but a more rigorous modeling approach is scope for future research.

We experimented with various indexes for multi market contact. The initial

[5] We are working on relaxing this assumption, focusing only on entry between deregulation (1990) and 2005. This would allow us to use pre-deregulation market structure as an exogenous determinant for the evolution of multi market contact during the free entry period. The exogeneity of the 1990's market structure can be justified with the strict regulation that froze market structure between 1936 and 1990, as described by Guiso et al. [2004]. The use of pre-existing firm's presence has been proposed first by Berry [1992], who exploited airlines' airport presence as a determinant of their route entry decisions.

option we considered was including all the links that banks had across the country, regarding a link as a LLS where both banks have at least one branch. This definition has at least two unrealistic assumptions. First, it implies that a price war in the north western Bardonecchia area, bordering France, could potentially cause retaliation in Siracusa, on the southern edge of Sicily, about 1150km far away. Mere geographical distance is not the only reason why this is unlikely to happen. The main justifications are the local nature of banking competition, as described in Degryse and Ongena [2005], and the hierarchical structure of Italian banks' decisions based on regional headquarters explained in Albareto et al. [2011]. The second controversial assumption is that banks could collude regardless of their market share in a LLS. This is also an unrealistic condition, because collusion is likely to be more effective in a concentrated oligopolistic market, as shown by Bernheim and Whinston [1990].

For these reasons we construct a LLS-bank specific multi market contact indicator that considers only the neighboring local markets where the local competitive fringe has less than 50% of total branches. We have run several robustness checks varying the distance of bordering LLSs as well as the market share of local banks. We find evidence that multi market contact matters less for greater distances or for higher share of branches of local players. In order to include the bordering markets in the multi market contact index we construct a measure of geographical distance between the main city councils of each LLS based on latitude and longitude coordinates. Given that the average surface of a LLS is about 400km^2, and assuming that the main city council is situated at the centre of the LLS, the mean distance between main municipalities is about 23km. We experiment including neighboring LLS that are up to 100km far away, finding a decline in the effect of the multi market index as distance grows.

We decide to exclude the largest metropolitan areas from the sample, as done in Cohen and Mazzeo [2007], to avoid having overlapping submarkets within a LLS. Hence, we drop the local markets above the 90th percentile of the population distribution. We have run some robustness checks and verified that our results are not sensitive to this particular cutoff point.

The estimation results are summarized in Table 4.3, 4.4, and 4.5. The first table shows some reduced form evidence from logit regressions for each of the 3 players. These results are then used as starting values for the structural estimation. The logit estimates show how the multi market contact index matters the most for closer distances and for lower market share of the competitive fringe. If on one hand the surface of the LLS doesn't seem to matter for banks' entry decision, on the other hand both population and value added have a positive and significant effect. The

employment rate appears to have a negative impact on entry.

The structural estimates in Table 4.4 confirm most of the reduced form results, but are partially different for the MMC indexes. Multiple links don't seem to matter at all for MPS, whereas they are increasingly significant for Sanpaolo as distance is reduced. Unicredit's profitability is significantly influenced by multi market contacts only for larger distances.

Finally, in Table 4.5 we look at different degrees of heterogeneity in the effect of multi market contact on banks' profitability, as in Ciliberto and Tamer [2009]. In this case we fix the distance of neighboring markets to 100km and just focus on Local Labor Systems where the share of local banks is less than 50%. We consider the case of homogeneous effect of multi market contact, heterogeneous across banks, and heterogeneous across bank-to-bank relationships. We find a positive and significant effect for the homogenous effect, a positive and significant effect for Unicredit and San Paolo in the first heterogeneous effect, and no statistically significant effect at the bank-to-bank relationship. This might be evidence that the three credit institutions don't value differently each of their competitors, but are rather affected by the sum of the contacts with their rivals.

Table 4.3: Reduced form logit results

Variable	**Distance<50km Local Share<50%**			**Distance<50km Local Share<40%**			**Distance<50km Local Share<60%**			**Distance<100km Local Share<50%**		
	UC	**SP**	**MPS**	**UC**	**SP**	**MPS**	**UC**	**SP**	**MPS**	**UC**	**SP**	**MPS**
Surface	0.936	-0.324	-0.498	0.919	-0.387	-0.761	0.904	-0.362	-0.489	0.955	-0.348	-0.722
	(0.664)	(0.575)	(0.478)	(0.669)	(0.572)	(0.470)	(0.661)	(0.571)	(0.478)	(0.668)	(0.574)	(0.470)
Population	0.583***	0.631***	0.475***	0.585***	0.634***	0.474***	0.560***	0.632***	0.470***	0.572***	0.626***	0.468***
	(0.079)	(0.069)	(0.051)	(0.078)	(0.069)	(0.050)	(0.079)	(0.069)	(0.051)	(0.078)	(0.069)	(0.050)
Value Added	1.859***	0.590*	0.596*	1.912***	0.617*	0.641*	1.813***	0.578*	0.603*	2.073***	0.684**	0.710**
	(0.401)	(0.343)	(0.337)	(0.396)	(0.344)	(0.333)	(0.401)	(0.343)	(0.339)	(0.401)	(0.343)	(0.335)
Employment Rate	-0.907**	-1.325***	-1.521***	-0.939***	-1.342***	-1.629***	-0.777**	-1.310***	-1.346***	-0.931**	-1.332***	-1.578***
	(0.374)	(0.344)	(0.309)	(0.368)	(0.344)	(0.305)	(0.377)	(0.346)	(0.313)	(0.373)	(0.344)	(0.307)
MMC Unicredit	0.514***	-	-	0.921***	-	-	0.315***	-	-	0.176***	-	-
	(0.083)	(-)	(-)	(0.167)	(-)	(-)	(0.046)	(-)	(-)	(0.028)	(-)	(-)
MMC Sanpaolo	-	0.153**	-	-	0.214**	-	-	0.084**	-	-	0.053**	-
	(-)	(0.061)	(-)	(-)	(0.109)	(-)	(-)	(0.038)	(-)	(-)	(0.021)	(-)
MMC MPS	-	-	0.294***	-	-	0.164	-	-	0.215***	-	-	0.056**
	(-)	(-)	(0.072)	(-)	(-)	(0.123)	(-)	(-)	(0.048)	(-)	(-)	(0.028)
Headquarter Unicredit	0.459	-	-	0.364	-	-	0.673	-	-	0.813	-	-
	(0.757)	(-)	(-)	(0.750)	(-)	(-)	(0.758)	(-)	(-)	(0.754)	(-)	(-)
Headquarter Sanpaolo	-	-1.461**	-	-	-1.403**	-	-	-1.405**	-	-	-1.315**	-
	(-)	(0.670)	(-)	(-)	(0.670)	(-)	(-)	(0.700)	(-)	(-)	(0.665)	(-)
Headquarter MPS	-	-	-5.077***	-	-	-4.861***	-	-	-4.579***	-	-	-4.626***
	(-)	(-)	(0.768)	(-)	(-)	(0.757)	(-)	(-)	(0.758)	(-)	(-)	(0.750)
Intercept	-0.831	4.116***	5.462***	-0.578	4.182***	6.100***	-1.600	4.031**	4.395***	-1.338	3.907**	5.577***
	(1.746)	(1.584)	(1.231)	(1.718)	(1.585)	(1.216)	(1.768)	(1.591)	(1.275)	(1.758)	(1.593)	(1.256)
Markets	617	617	617	617	617	617	617	617	617	617	617	617
R^2	0.379	0.313	0.303	0.372	0.310	0.285	0.388	0.311	0.310	0.374	0.313	0.288
LR χ^2	293.64	255.63	259.11	288.31	253.21	243.79	300.28	254.28	263.35	289.52	255.69	245.98
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Standard errors are in brackets. * is 10% significance level, ** is 5% significance level, *** is 1% significance level.

Table 4.4: Structural estimates varying distance from the market and share of local banks

Variable	**Distance<50km Local Share<50%**	**Distance<50km Local Share<40%**	**Distance<50km Local Share<60%**	**Distance<100km Local Share<50%**
Surface	-0.533*	-0.571*	-0.414	-0.408
	(0.299)	(0.301)	(0.308)	(0.289)
Population	0.513***	0.517***	0.528***	0.495***
	(0.032)	(0.031)	(0.033)	(0.034)
Value Added	1.001***	1.042***	0.979***	0.956***
	(0.188)	(0.186)	(0.192)	(0.181)
Employment Rate	-1.548***	-1.520***	-1.571***	-1.428***
	(0.186)	(0.187)	(0.192)	(0.182)
MMC Unicredit	-0.142	0.490	-0.097	0.798**
	(0.389)	(0.569)	(0.406)	(0.346)
MMC Sanpaolo	1.652***	3.037***	0.302	0.785**
	(0.473)	(0.881)	(0.336)	(0.333)
MMC MPS	-0.140	-0.489	-0.498*	-0.282
	(0.288)	(0.364)	(0.262)	(0.244)
Headquarter Unicredit	-0.838***	-0.946***	-0.997***	-0.948**
	(0.391)	(0.382)	(0.415)	(0.377)
Headquarter Sanpaolo	-1.635***	-1.562***	-1.498***	-1.349***
	(0.365)	(0.366)	(0.383)	(0.371)
Headquarter MPS	-4.095***	-4.010***	-4.001***	-3.358***
	(0.559)	(0.551)	(0.578)	(0.555)
Intercept	5.081***	4.904***	5.225***	4.425***
	(0.811)	(0.819)	(0.841)	(0.800)
μ	-4.904***	-4.911***	-4.903***	-4.907
	(0.202)	(0.202)	(0.202)	(0.202)
σ	4.861***	4.857***	4.863***	4.859***
	(0.679)	(0.673)	(0.682)	(0.681)
Markets	617	617	617	617

Note: Standard errors are in brackets. * is 10% significance level, ** is 5% significance level, *** is 1% significance level.

Table 4.5: Structural estimates varying heterogeneity in the effect of multi market contact

Variable	Homogeneous MMC	Bank MMC	Bank-to-Bank MMC
Surface	0.829	0.870	0.888
	(0.903)	(0.898)	(0.897)
Population	2.012***	2.003***	1.994***
	(0.096)	(0.095)	(0.114)
Value Added	0.537***	0.532***	0.533***
	(0.060)	(0.060)	(0.061)
Employment Rate	-0.615***	-0.615***	-0.612***
	(0.061)	(0.061)	(0.063)
MMC	0.324***	-	-
	(0.089)		
MMC Unicredit	-	0.517*	-
		(0.356)	
MMC Sanpaolo	-	0.609*	-
		(0.348)	
MMC MPS	-	0.0003	-
		(0.274)	
MMC Unicredit-Sanpaolo	-	-	0.0005
			(2.288)
MMC Unicredit-MPS	-	-	1.627
			(3.373)
MMC Sanpaolo-Unicredit	-	-	0.883
			(2.212)
MMC Sanpaolo-MPS	-	-	0.002
			(3.433)
MMC MPS-Unicredit	-	-	0.001
			(6.531)
MMC MPS-Sanpaolo	-	-	0.001
			(7.036)
Headquarter Unicredit	0.049	-0.016	-0.037
	(0.179)	(0.224)	(0.247)
Headquarter Sanpaolo	-0.177	-0.281	-0.227
	(0.167)	(0.215)	(0.238)
Headquarter MPS	-1.403***	-1.173***	-1.185***
	(0.226)	(0.260)	(0.300)
Intercept	0.138***	0.137***	0.135***
	(0.027)	(0.027)	(0.029)
μ	-0.035	-0.035	-0.035
	(0.058)	(0.058)	(0.059)
σ	0.462***	0.463***	0.463***
	(0.015)	(0.015)	(0.015)
Markets	617	617	617

Note: Standard errors are in brackets. In this specification we always keep distance less than 100km and share of local players less than 50%. * is 10% significance level, ** is 5% significance level, *** is 1% significance level. The coefficients for "Bank MMC" are slightly different in magnitude from the last column in Table 4.4 due to a small difference in the scaling of some variables.

4.6 Conclusion

This paper investigates the effect of multi market contact on entry decisions of the main Italian banks. It extends a well established static model of market structure with incomplete information, developed by Seim [2006], allowing for endogenous multi market contact creation and heterogeneous bank profitability. We focus on the 2005 market structure for the three main credit institutions in Italy: Unicredit, Intesa San Paolo, and Monte dei Paschi di Siena. We also use a novel definition of geographical markets in Italian banking, that are the local labor systems (LLS) instead of provinces. These local markets are comparable in size to the Labor Market Areas in the US, used for most of the entry literature in the US, and are defined according to integrated economic areas and commuting patterns between councils. For these reasons, the LLS are likely to represent more accurately the area where banks compete.

We analyze how the number of contacts that a multi market bank has with its potential rivals affects its entry decision in a specific market. We find that multiple links in the neighboring markets enhance banks' profitability, with a stronger effect when the local competitive fringe has a lower market share and when the number of neighboring markets is bigger. We also allow for different degrees of heterogeneity in the effect of multi market contact on profitability across banks, and still find a positive impact, apart from the bank-to-bank relationship level. Therefore, our results suggest that banks find it profitable to establish contacts between each other, which might facilitate implicit collusion as in Bernheim and Whinston [1990].

4.7 Appendix

The derivation of the polynomial starts from equation 9, expressed here for the case of 3 players:

$$\begin{aligned}\Omega_m &= p_{1m} + p_{2m} + p_{3m} \\ &= \frac{\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]}{1+\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]} + \frac{\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]}{1+\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]} + \frac{\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]}{1+\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]}.\end{aligned} \tag{4.12}$$

This expression can be developed as:

$$\begin{aligned}\Omega_m \quad * \quad & \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]\big] \\ = \quad & \big[\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]\big] \\ + \quad & \big[\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]\big] \\ + \quad & \big[\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)]\big] * \big[1+\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)]\big],\end{aligned} \tag{4.13}$$

where the LHS of equation 13 becomes:

$$\begin{aligned}& \Omega_m + \Omega_m \exp(\xi^m)\Big[\exp(\tilde{\Pi}_1^m) + \exp(\tilde{\Pi}_2^m) + \exp(\tilde{\Pi}_3^m)\Big] \\ + \quad & \Omega_m \exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] + [\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_3^m)] + [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\ + \quad & \Omega_m \exp(\xi^m)^3\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big\},\end{aligned} \tag{4.14}$$

and the RHS of equation 13 becomes:

$$\begin{aligned}& \Big\{\exp(\xi^m)[\exp(\tilde{\Pi}_1^m)] + \exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] + [\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\ & + \exp(\xi^m)^3[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big\} \\ + \quad & \Big\{\exp(\xi^m)[\exp(\tilde{\Pi}_2^m)] + \exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] + [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\ & + \exp(\xi^m)^3[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big\} \\ + \quad & \Big\{\exp(\xi^m)[\exp(\tilde{\Pi}_3^m)] + \exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_3^m)] + [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\ & + \exp(\xi^m)^3[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big\}.\end{aligned} \tag{4.15}$$

The RHS can be further simplified to:

$$
\begin{aligned}
& \exp(\xi^m)\Big[\exp(\tilde{\Pi}_1^m) + \exp(\tilde{\Pi}_2^m) + \exp(\tilde{\Pi}_3^m)\Big] \\
+ \quad & 2\exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] + [\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_3^m)] + [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\
+ \quad & 3\exp(\xi^m)^3\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big].
\end{aligned}
\tag{4.16}
$$

Bringing now back together LHS and RHS we get:

$$
\begin{aligned}
& (3-\Omega_m)\exp(\xi^m)^3\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\
+ \quad & (2-\Omega_m)\exp(\xi^m)^2\Big[[\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_2^m)] + [\exp(\tilde{\Pi}_1^m)] * [\exp(\tilde{\Pi}_3^m)] + [\exp(\tilde{\Pi}_2^m)] * [\exp(\tilde{\Pi}_3^m)]\Big] \\
+ \quad & (1-\Omega_m)\exp(\xi^m)\Big[\exp(\tilde{\Pi}_1^m) + \exp(\tilde{\Pi}_2^m) + \exp(\tilde{\Pi}_3^m)\Big] - \Omega_m = 0,
\end{aligned}
\tag{4.17}
$$

which is equivalent to equation 10 in Section 3.

Chapter 5

Conclusion

This thesis is formed by three coherently connected papers that investigate the relationship between asymmetric information, imperfect competition, and market structure in the Italian banking sector.

The first paper measures the welfare costs of asymmetric information and imperfect competition in the market for small business credit lines. First, it presents reduced form evidence of adverse selection and moral hazard, conducting a positive correlation test à la Chiappori and Salanié [2000]. Based on these results, it develops a structural model of firms' demand for credit, loan use and default, and of banks' pricing. It provides a simplified Monte Carlo simulation of the model, showing that imperfect competition can mitigate the effect of adverse selection. The estimates of the structural model report presence of adverse selection, but not of moral hazard. Finally, two counterfactual policy experiments are constructed. The first one recovers the welfare costs of asymmetric information, the second one predicts the consequences of a credit crunch on a credit market with asymmetric information and imperfect competition.

This paper could be extended in several dimensions. A first direction could entail enriching the theoretical foundations of the baseline model to separately identify asymmetric information and risk aversion. So far the paper assumes risk neutrality on both the borrower's and the lender's side, as in Stiglitz and Weiss [1981]. However, following the example of Cohen and Einav [2007], expected utility theory could help to recover risk aversion for borrowers. Along these lines, it would be interesting also to estimate a risk aversion parameter for lenders, especially in the context of a financial crisis, possibly allowing for risk aversion heterogeneity among banks. The supply side of this paper is another part that could be extended, or at least modeled differently. Currently we rely on the assumption of posted prices

for new borrowers, even though we don't observe these prices directly in the data. It could be also reasonable to think that interest rates charged by lenders are the results of a bargaining process with borrowers, in the presence of asymmetric information. It would be also interesting to extend the welfare analysis considering other counterfactual policy experiments, such as reducing the extent of competition simulating a bank merger.

The second paper builds on the first paper's model, analyzing the effect of asymmetric information on the banking market structure. It begins providing descriptives and reduced form evidence of how experienced incumbent banks have an informational advantage over potential entrants, which allows them to have a more creditworthy pool of borrowers. Moreover, it is showed that the longer experience of incumbent banks plays a significant role in deterring new banks from entering. Given these results, the paper develops a dynamic structural model of banks' entry, exit and investment through branching, which allows them to gain experience and learn about their borrowers. Using the notion of oblivious equilibrium developed by Weintraub et al. [2008b] gives a tractable solution to the computational problem of many players and a nonstationary environment.

This paper sets the basis for various possible extensions. First, it allows for a relatively reduced form learning process based on the definition of experience. This could be enriched assuming that lenders learn through time either the firm-specific default probability, or a market-bank specific correlation coefficient that identifies asymmetric information. Second, it would be important to have a better understanding of the dynamic evolution of the relationships between borrowers and lenders. How does a lender acquire soft information about a borrower? How is this reflected into the pricing strategy? Do search and switching costs increase for borrowers as the relationship with a lender evolves? Petersen and Rajan [1995] provide theoretical grounds to answer to these questions, looking at the effect of credit market competition on lending relationships. Third, this paper focuses primarily on entry barriers from informational asymmetries in the credit market. It could be the case however that there are other sources of entry barriers, such as scale economies, as banks might have a different cost structure depending on their size, or as scope economies, as banks might have heterogeneous costs depending on their area of lending specialization.

The last paper investigates the relationship between multi market contacts between banks and market structure. It concentrates on the main Italian credit institutions, and develops a static structural model of market structure based on Seim [2006], with endogenous multi market contact creation. The main finding of

the paper is that multiple links with rivals in neighboring markets have a positive and significant effect on a banks' decision to enter into a market. This effect is stronger for closer markets and for markets where the main banks have a stronger presence, suggesting that multi market contact might facilitate collusion as in Bernheim and Whinston [1990].

This work could also be extended in various ways. One of its main limitations is the static approach, as entry and the creation of contacts across markets are inherently dynamic decisions. Hence, using a dynamic model would be a natural extension. Another improvement could concern the specification of the profit function. So far we have assumed a reduced form profit function, for lack of data on prices and quantities. Having access to additional data would allow us to model demand and pricing, and to construct a profit function based on these fundamentals. Within such a structural model we could allow banks to collude on prices, as Ciliberto and Williams [2013] did for the airline industry, where they identified conduct parameters using multi market contacts. We could then allow banks to take entry and exit decisions based on this structural model with collusion and endogenous multi market contact creation.

Bibliography

J. H. Abbring, P. A. Chiappori, J. J. Heckman, and J. Pinquet. Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association*, 1:512–521, 2003.

D.A. Ackerberg and M. Botticini. Endogenous matching and the empirical determinants of contract form. *Journal of Political Economy*, 110(3):564–591, 2002.

V. Aguirregabiria and C. Y. Ho. A dynamic oligopoly game of the us airline industry: Estimation and policy experiments. *Journal of Econometrics*, 168(1):156–173, 2012.

G. Akerlof. The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

G. Albareto, M. Benvenuti, S. Mocetti, M. Pagnini, and P. Rossi. The organization of lending and the use of credit scoring techniques in italian banks. *The Capco Institute Journal of Financial Transformation*, 32:143–157, 2011.

E. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 21:589–609, 1968.

E. Altman, G. Marco, and F. Varetto. Comporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the italian experience). *Journal of Banking and Finance*, 18:505–529, 1994.

K. Arrow. The economic implications of learning by doing. *Review of Economic Studies*, 29(3):155–173, 1962.

K. Arrow. Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5):941–973, 1963.

P. Bajari, C. L. Benkard, and J. Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

A. Banerjee and A. Newman. Occupational choice and the process of development. *Journal of Political Economy*, 101(2):274–298, 1993.

C. L. Benkard. Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90(4):1034–1054, 2000.

C. L. Benkard. A dynamic analysis of the market for wide-bodied commercial aircraft. *Review of Economic Studies*, 71:581–611, 2004.

C. L. Benkard, P. Jeziorski, and G. Y. Weintraub. Oblivious equilibrium for concentrated industries. NBER Working Paper 19307, 2013.

A. Bernanke and M. Gertler. Financial fragility and economic performance. *Quarterly Journal of Economics*, 105(1):87–114, 1990.

B.D. Bernheim and M.D. Whinston. Multimarket contact and collusive behavior. *RAND Journal of Economics*, 21(1):1–26, 1990.

S. Berry. Estimation of a model of entry in the airline industry. *Econometrica*, 60 (4):889–917, 1992.

S. Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, 1994.

S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

S. Berry, J. Levinsohn, and A. Pakes. Differentiated products demand systems from a combination of micro and macro data: The new vehicle market. *Journal of Political Economy*, 112(1):68–104, 2004.

D. Besanko, U. Doraszelski, Y. Kryukov, and M. Satterthwaite. Learning-by-doing, organizational forgetting, and industry dynamics. *Econometrica*, 78(2):453–508, 2010.

M. Bofondi and G. Gobbi. Informational barriers to entry into credit markets. *Review of Finance*, 10:39–67, 2006.

T.F. Bresnahan and P. Reiss. Entry and competition in concentrated markets. *Journal of Political Economy*, 99(5):977–1009, 1991.

T. Broecker. Credit-worthiness tests and interbank competition. *Econometrica*, 58 (2):429–452, 1990.

M.C. Byford and J.S. Gans. Collusion at the extensive margin. NBER Working Paper 20163, 2014.

L.M.B. Cabral and M.H. Riordan. The learning curve, market dominance, and predatory pricing. *Econometrica*, 62(5):1115–1140, 1994.

P. A. Chiappori and B. Salanié. Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78, 2000.

M. Chicu and N. L. Ziebarth. Multi-market contact and competition: Evidence from the depression-era portland cement industry. *International Journal of Industrial Organization*, xxx(xxx):xxx, 2013.

L. Ciari and R. De Bonis. Entry decisions after deregulation: the role of incumbents' market power. Mofir Working Paper, 2011.

L. Ciari and N. Pavanini. Market structure and multi market contact. Working Paper, 2013.

F. Ciliberto and E. Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.

F. Ciliberto and J.W. Williams. Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry. Working Paper, 2013.

P. Coccorese and A. Pellecchia. Multimarket contact and profitability in banking: Evidence from italy. *Journal of Financial Services Research*, 35:245–271, 2009.

A. Cohen and L. Einav. Estimating risk preferences from deductible choice. *American Economic Review*, 97(3):745–788, 2007.

A. M. Cohen and M. Mazzeo. Market structure and competition among retail depository institutions. *The Review of Economics and Statistics*, 89(1):60–74, 2007.

G.S. Crawford, N. Pavanini, and F. Schivardi. Asymmetric information and imperfect competition in the loan market. CAGE - University of Warwick Working Paper 167/2013, 2013.

P. Dasgupta and J. Stiglitz. Learning-by-doing, market structure and industrial trade policies. *Oxford Economic Papers*, 40:246–268, 1988.

R. De Bonis and A. Ferrando. The italian banking structure in the 1990s: Testing the multimarket contact hypothesis. *Economic Notes*, 29(2):215–241, 2000.

R. de Elejalde. Local entry decisions in the us banking industry. Working Paper Universidad de Chile, 2012.

H. Degryse and S. Ongena. Distance, lending relationships, and competition. *Journal of Finance*, LX(1):231–266, 2005.

G. Dell'Ariccia. Learning by lending, competition, and screening incentives in the banking industry. IMF - Research Department Working Paper, 2000.

G. Dell'Ariccia. Asymmetric information and the structure of the banking industry. *European Economics Review*, 45:1957–1980, 2001.

G. Dell'Ariccia, E. Friedman, and R. Marquez. Adverse selection as a barrier to entry in the banking industry. *RAND Journal of Economics*, 30(3):515–534, 1999.

D. DeMeza and D.C. Webb. Too much investment: A problem of asymmetric information. *The Quarterly Journal of Economics*, 102(2):281–292, 1987.

E. Detragiache, P. Garella, and L. Guiso. Multiple versus single banking relationships: Theory and evidence. *The Journal of Finance*, LV(3):1133–1161, 2000.

C.D. Edwards. Conglomerate bigness as a source of power. *Business Concentration and Price Policy*, pages 331–359, 1955. NBER Conference Report, Princeton: Princeton University Press.

L. Einav, A. Finkelstein, and J. Levin. Beyond testing: Empirical models of insurance markets. *Annual Review of Economics*, 2:311–336, 2010a.

L. Einav, A. Finkelstein, and P. Schrimpf. Optimal mandates and the welfare cost of asymmetric information: Evidence from the u.k. annuity market. *Econometrica*, 78(3):1031–1092, 2010b.

L. Einav, M. Jenkins, and J. Levin. Contract pricing in consumer credit markets. *Econometrica*, 80(4):1387432, 2012.

R. Ericson and A. Pakes. Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62:53–82, 1995.

W.N. Evans and I.N. Kessides. Living by the golden rule: Multimarket contact in the u.s. airline industry. *Quarterly Journal of Economics*, 109:341–366, 1994.

R.M. Feinberg. Mutual forbearance as an extension of oligopoly theory. *Journal of Economics and Business*, 36:243–249, 1984.

R. Felici and R. Pagnini. Distance, bank heterogeneity and entry in local banking markets. *Journal of Industrial Economics*, 53(3):500–534, 2008.

L. Fuentelsaz and J. Gomez. Multipoint competition, strategic similarity, and entry into geographic markets. *Strategic Management Journal*, 27:477–499, 2006.

W. G. Gale. Federal lending and the market for credit. *Journal of Public Economics*, 42(2):177–193, 1990.

J. Gerakos and C. Syverson. Competition in the audit market: Policy implications. Working Paper, 2014.

G. Gobbi and F. Lotti. Entry decisions and adverse selection: An empirical analysis of local credit markets. *Journal of Financial Services Research*, 26(3):225–244, 2004.

A. Goolsbee and A. Petrin. The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica*, 72(2):351–381, 2004.

L. Guiso, P. Sapienza, and L. Zingales. Does local financial development matter? *Quarterly Journal of Economics*, 119(3):929–969, 2004.

L. Guiso, P. Sapienza, and L. Zingales. The cost of banking regulation. CEPR Discussion Papers 5864, 2007.

L. Guiso, L. Pistaferri, and F. Schivardi. Credit within the firm. *The Review of Economic Studies*, 80(1):211–247, 2013.

B. Handel. Adverse selection and switching costs in health insurance markets: When nudging hurts. Forthcoming *American Economic Review*, 2013.

J. A. Hausman and W. E. Taylor. Panel data and unobservable individual effects. *Econometrica*, 49(6):1377–1398, 1981.

J.J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1): 153–161, 1979.

R. G. Hubbard. Capital market imperfections and investment. *Journal of Economic Literature*, 36(1):193–225, 1998.

K. Hughes and C. Oughton. Diversification, multimarket contact and profitability. *Economica*, 60:203–224, 1993.

I. Jans and D.I. Rosenbaum. Multimarket contact and pricing: Evidence from the u.s. cement industry. *International Journal of Industrial Organization*, 15:391–412, 1993.

D. Karlan and J. Zinman. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica*, 77(6):1993–2008, 2009.

J. Lustig. Measuring welfare losses from adverse selection and imperfect competition in privatized mediare. Boston University Working Paper, 2011.

N. G. Mankiw. The allocation of credit and financial collapse. *Quarterly Journal of Economics*, 101(3):455–470, 1986.

R. Marquez. Competition, adverse selection, and information dispersion in the banking industry. *Review of Financial Studies*, 15(3):901–926, 2002.

M. J. Mazzeo. Product choice and oligopoly market structure. *RAND Journal of Economics*, 33(2):1–22, 2002.

J. Molnar, R. Violi, and X. Zhou. Multimarket contact in italian retail banking: Competition and welfare. *International Journal of Industrial Organization*, 31(5): 368–381, 2011.

D. Mookherjee and D. Ray. Contractual structure and wealth accumulation. *American Economic Review*, 92(4):818–849, 2002.

A. Nevo. Mergers with differentiated products: the case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 31(3):395–421, 2000.

A. Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, 2001.

M. Pagano and T. Jappelli. Information sharing in credit markets. *Journal of Finance*, XLVIII(5):1693–1718, 1993.

A. Pakes and P. McGuire. Computing markov-perfect nash equilibria: Numerical implications of a dynamic differentiated product model. *The RAND Journal of Economics*, 25(4):555–589, 1994.

F. Panetta, F. Schivardi, and M. Shum. Do mergers improve information? evidence from the loan market. *Journal of Money, Credit and Banking*, 41(4):673–709, 2009.

N. Pavanini and F. Schivardi. Dynamic entry and exit with learning by branching. University of Warwick Working Paper, 2013.

H. Perez-Saiz. Building new plants or entering by acquisition? firm heterogeneity, identification and entry barriers. Revise and Resubmit at *RAND Journal of Economics*, 2013.

M. A. Petersen and R. G. Rajan. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110:403–444, 1995.

M. A. Petersen and R. G. Rajan. Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570, 2002.

A. Polsi. *L'articolazione territoriale del sistema bancario italiano fra scelte di mercato e intervento delle autorita' monetarie.* Il Mulino, Bologna, 2000. in G. Conti and S. La Francesca *Banche e reti di banche nell'Italia Postunitaria.*

S. Qi. The impact of advertising regulation on industry: The cigarette advertising ban of 1971. *The RAND Journal of Economics*, 44(2):215–248, 2013.

M. Rothschild and J.E. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90(4):630–649, 1976.

S.P. Ryan. The costs of environmental regulation in a concentrated industry. *Econometrica*, 80(3):1019–1061, 2012.

K. Seim. An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics*, 37(3):619–640, 2006.

S. Shaffer. The winner's curse in banking. *Journal of Financial Intermediation*, 7: 359–392, 1998.

G. Spagnolo. On interdependent supergames: Multimarket contact, concavity, and collusion. *Journal of Economic Theory*, 89:127–139, 1999.

M. A. Spence. The learning curve and competition. *Bell Journal of Economics*, 12 (1):49–70, 1981.

A. Starc. Insurer pricing and consumer welfare: Evidence from medigap. RAND Journal of Economics forthcoming, 2013.

J. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410, 1981.

O. Toivanen and M. Waterson. Market structure and entry: Where's the beef? *RAND Journal of Economics*, 36(3):680–699, 2005.

K. E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2009.

K. E. Train and C. Winston. Vehicle choice behavior and the declining market share of u.s. automakers. *International Economic Review*, 48(4):1469–1496, 2007.

G. Y. Weintraub, C. L. Benkard, P. Jeziorski, and B. Van Roy. Nonstationary oblivious equilibrium. Working Paper, Columbia University, 2008a.

G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411, 2008b.

G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Computational methods for oblivious equilibrium. *Operations Research*, 58(3):1247–1265, 2010.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2002.

D. Y. Xu. A structural empirical model of r&d, firm heterogeneity, and industry evolution. Working Paper, New York University, 2008.