Review of Finance (2006) 10: 39–67
DOI 10.1007/s10679-006-6978-2

Informational Barriers to Entry into Credit Markets *

MARCELLO BOFONDI and GIORGIO GOBBI
Bank of Italy

Abstract. Economic theory suggests that asymmetric information between incumbents and entrants can generate barriers to entry into credit markets. Incumbents have superior information about their own customers and the overall economic conditions of the local credit market. This implies that entrants are likely to experience higher loan default rates than the incumbents. We test these theoretical predictions using a unique database of 7,275 observations on 729 individual banks' lending in 95 Italian local markets. We find that informational asymmetries play a significant role in explaining entrants' loan default rates. The default rate is significantly higher for those banks that entered local markets without opening a branch, suggesting that having a branch on site may help to reduce the informational disadvantage. We also uncover a positive correlation between banks' loan default rates in individual local markets and the number of banks lending in that market. We argue that these informational barriers can help to explain why entry into many local credit markets by domestic and foreign banks was slow, even after substantial deregulation.

1. Introduction

Geographical restrictions and legal barriers to entry into the banking industry have been progressively relaxed since the late 1970s in the United States, in Europe – both at country level and at European Union level – and in many other countries.[1] The lifting of regulatory constraints has been followed by an increase in competition and, according to a large number of studies, by substantial gains in terms of consumer welfare.[2] Nonetheless, US banks have been slower than expec-

* The views expressed in this article are those of the authors and do not involve the responsibility of the Bank of Italy. The authors thank Franklin Allen, Dario Focarelli, Andrea Generale, Luigi Guiso, Francesca Lotti, Marco Pagano, Alberto Franco Pozzolo, Paola Sapienza, Alessandro Secchi, two anonymous referees and seminar participants at the Bank of Italy, the Federal Reserve of Chicago, the 2003 BIS Workshop on Applied Banking Research, the 2003 EARIE Conference, the First Banca d'Italia/CEPR Conference on Money, Banking and Finance, the 2004 FIRS Conference on Banking, Insurance and Intermediation and the 2004 EEA Meeting for their comments. The usual disclaimer applies to all of them.

[1] See Berger et al. (1995) and Vives (1991) for a survey of the major regulatory changes in the US and the European Union respectively.

[2] Hannan and Prager (1998) and Angelini and Cetorelli (2003).

ted to move into new markets and cross-border banking in Europe is still limited, particularly in the retail market.[3]

The experience so far accumulated is broadly consistent with the predictions of a strand of theoretical research that emphasizes the role of asymmetric information in shaping the structure of credit markets. According to this literature, entrants to credit markets are at a disadvantage because incumbents have proprietary information about borrowers' risk characteristics and are consequently able to cream the market and leave the newcomers to cope with riskier customers.

The question of whether informational entry barriers are a key feature of credit markets is very important from a policy perspective. If they are, geographical segmentation is likely to be a persistent feature of credit markets, regardless of deregulation or technological advances.[4] Allowing free entry into the banking industry may not be sufficient to help local financial development, which has been proved to play a fundamental role in supporting and fostering economic growth (King and Levine, 1993; Guiso et al., 2004). Further, even if entry barriers are not too high to block entry, the additional risk that entrant banks have to bear increases their probability of default, and may thereby reduce the stability of the banking system.

We empirically address these issues using microeconomic data on Italian local credit markets in the early 1990s. A distinctive feature of the Italian economy is the large number of small and medium-sized firms, which tend to cluster in tight local business communities. Regulatory changes that were introduced in the banking sector as of the late 1980s allowed a substantial increase in the number of banks operating in each local market. Thanks to access detailed bank-level data, we are able to compare individual banks' loan default rates in each local market and to relate them to measures of the information available both to incumbents and to entrants.

There are two main arguments supporting the view that asymmetric information between banks works as a barrier to entry into credit markets. The first, known as the winner's curse hypothesis, derives from the fact that when a bank assesses the creditworthiness of potential borrowers it does not know whether they have already been rejected by the other banks in the market (Broecker, 1990; Riordan, 1993). If the tests run by different banks are not perfectly correlated, there is a positive probability that an applicant's creditworthiness will be assessed differently by different lenders. It follows that the average quality of the pool of successful borrowers declines as the number of banks in the market rises (and consequently the expected losses from bad loans increase). Adverse selection is greater for entrants because their pool of applicants includes the would-be borrowers who were previously

[3] As documented in Buch (2002) and Danthine et al. (1999).

[4] Petersen and Rajan (2002) document that distance still plays a major role in the provision of credit to small firms, even though its importance is diminishing with the growth of remote banking facilities. Degryse and Ongena (2005) report comprehensive evidence of spatial price discrimination in bank lending caused by transportation costs.

rejected by the mature banks in the market (Shaffer, 1998). The second argument relies on the fact that a substantial amount of the information used by banks to screen loan applicants and monitor borrowers is generated through repeated interaction with their customers and the local business community (Dell'Ariccia et al., 1999; Marquez, 2002). It follows that the creditworthiness tests of incumbents are likely to be more accurate than those run by entrants.

Compared with the abundance of theoretical papers on the subject, empirical work has been rather limited. To the best of our knowledge, Shaffer (1998) is the only study that empirically tests the winner's curse in banking. He finds that net chargeoff rates of *de novo* banks are significantly higher than those recorded by mature banks. Shaffer also estimates a cross-sectional model using data from mature banks, each operating in a single geographic market. He finds a strong positive correlation between gross chargeoff rates and the total number of banks in each market area.

This paper tests whether incumbents have superior information than entrants by investigating the variation in individual banks' loan default rates across local markets. Our identification hypothesis is that, after controlling for bank and market characteristics, systematic differences in loan default rates reflect systematic differences in the pools of borrowers faced by each lender.

As predicted by theory, we find evidence that, after controlling for the level of information on local market conditions, entrants are more exposed to bad loans than incumbents. Our findings are consistent with the hypothesis that incumbents select a less risky pool of borrowers. We do not have information on interest rates, hence we can not infer whether entrants are able to charge adequate risk premia. Nevertheless the fact that they can not replicate the incumbents' portfolio shows that they have a competitive disadvantage.

Subsequently, we distinguish between two different ways in which a bank can enter a local market. The first consists in the bank's branches or headquarter granting loans from outside the local market. The second way is for the bank to open a branch. We argue that the two types of entry differ substantially in terms of the information gap between entrants and incumbents. Having on site branches allows banks to monitor borrowers more rigorously and to acquire more detailed knowledge of the local economy. We find that banks that enter by opening a branch experience lower default rates than those that enter without opening a branch.

Consistently with the result obtained by Shaffer, we also find a positive correlation between banks' loan default rates in individual local markets and the number of banks lending in that market. We show that the disadvantage of entrants is smaller in markets where borrowers' turnover is higher. Lastly, we investigate the effects of entry on incumbents' default rates. We find that the quality of incumbents' loan portfolios in a given local market is positively correlated with the number of entrants in that market.

The remainder of this paper is organized as follows. In Section 2 we present our data and explain how the relevant variables have been constructed. Section

3 presents the empirical model used to investigate the relationship between entry and loan default rates and some relevant extensions. The robustness of the results with respect to potential endogeneity problems is tested in Section 4. In Section 5 we explore the relationship between bank and market characteristics and loan default rates. Section 6 is devoted to discuss an alternative interpretation of the main results, namely that the positive relationship between entry and loan default rates may be due to organizational differences across banks. In Section 7 we analyze the effect of entry on incumbents' loan default rates. Section 8 concludes.

2. Data

2.1. ENTRY INTO ITALIAN LOCAL CREDIT MARKETS

We use a data set on bank lending in Italy over the period from 1986 to 1996. It contains detailed information on individual banks' credit volumes and default rates for 95 local markets. Local markets are defined as provinces, i.e., the smallest local governments for which a satisfactory set of economic statistics exists, which are very similar to the U.S. counties in terms of size. Although provinces do not necessarily correspond to well-defined business communities, they represent good approximations for local credit markets. On average, 80 per cent of borrowing by residents in a given province is from bank branches in the same province.[5]

The main purpose of this paper is to test whether entrants to a local credit market are systematically subject to higher loan default rates than incumbents. The loan default rate is an ex-post measure of risk and thus depends on the states of nature that trigger defaults. In normal times defaults tend to be driven mainly by idiosyncratic shocks affecting individual borrowers, so that empirically it can be very difficult to test differences in banks' choices. To isolate the effects of the latter, we choose to measure default rates in the aftermath of an economic downturn that represents a common shock to all borrowers. The sequence of events that we need for our empirical strategy is as follows. At time T_0, given some initial conditions for banks and local markets, reforms are introduced that lower the barriers to entry into local markets. Between time T_0 and T_1, thanks to the deregulation, banks can choose to expand their networks by entering new markets. Finally, at T_2 a common shock occurs and the proportion of insolvent loans in banks' portfolios is revealed.

The Italian case fits nicely into this pattern. Beginning in the 1980s the Italian banking system underwent a series of reforms aimed at increasing competition in the market. Italian credit authorities increasingly liberalized branching and eased the geographical restrictions on lending, thereby lowering the legal barriers to entry into local markets. From the late 1970s onwards the opening of new branches had been regulated by "branch distribution plans" which were issued every four years.

[5] In previous research on the Italian banking industry local markets have been identified with provinces by Bonaccorsi di Patti and Gobbi (2001), Sapienza (2002) and Focarelli and Panetta (2004).

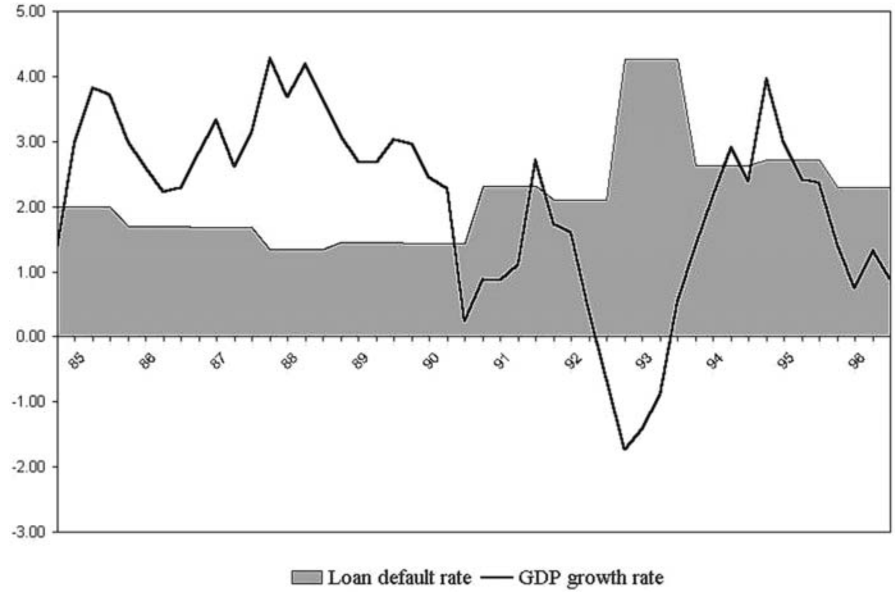

Figure 1. Bad Loans of Italian Banks. The graph plots the loan default rate of Italian banks and the Italian GDP growth rate for the period 1985–1996. The loan default rate is defined as new bad loans as a percentage of the stock of prior-year loans not in default: annual data. The GDP growth rate is computed at constant prices with respect to the corresponding quarter of the preceding year: quarterly data.

The last distribution plan was issued in 1986; as of March 1990, the establishment of new branches was completely liberalized. This led to an unprecedented increase in the number of branches: from 13,136 in 1985 to 19,786 in 1992. The phenomenon was nationwide: on average, the number of branches per province grew by 49.3 per cent with a standard deviation of 23.1. From 1985 to 1991 the Italian economy enjoyed a period of growth and the credit-to-GDP ratio also rose significantly. GDP growth started to slow down at the end of 1991 and bottomed out during the first quarter of 1993. The long and severe recession proved to be a hard test for loans granted in the previous years: the overall loan default rate rose to 4.2 per cent from an average of 1.5 per cent in the previous years (Figure 1). Several banks suffered huge losses and the profitability of the Italian banking system was depressed for several years to come.

2.2. SOURCES AND DATA DESCRIPTION

We draw our bank data from two sources: banks' balance sheet and income statement information are taken from the Supervisory Reports to the Bank of Italy; default rates and other credit variables regarding local credit markets from the Italian Central Credit Register (CCR).[6] The CCR is a department of the Bank of

[6] A description of the Italian CCR is contained in Miller (2003).

Italy that collects data on borrowers from their lending banks. Reporting banks file detailed information for each borrower with total loans and credit lines of more than 75,000 euros. Banks are requested to report smaller exposures only in the case of the borrower's default. Bad loans are defined on a customer basis and therefore include all the outstanding credit extended by a bank to a borrower considered insolvent.

According to Italian regulations, banks are asked to classify outstanding loans to the borrowers that are not expected to meet their obligations as bad loans. This allows banks a certain amount of discretion in judging whether a loan is bad or not. We use an "adjusted" definition of bad loans provided by the CCR. In the case of a single bank relationship this definition coincides with that of bad loans, namely it covers all the loans extended to the insolvent borrowers. Loans extended to borrowers with multiple bank relationships are all classified as ("adjusted") bad loans when: (i) the borrower is reported as insolvent by a bank that accounts for 70 per cent or more of the exposure to the banking system; (ii) the borrower is reported as insolvent by two or more banks that account for at least 10 per cent of the total exposure to the banking system.

Our sample has 7,275 observations referring to 729 banks, representing virtually the entire population of Italian credit institutions, which in that period included commercial banks, thrift institutions, special credit institutions, cooperative and community banks. We excluded de novo banks since they may exhibit atypical performance during their first few years (DeYoung and Hasan, 1998). Mergers and acquisitions that took place between 1986 and 1996 were considered as if they had occurred in 1986; this implies that entry by M&A is not considered in our analysis. The rationale for this choice is twofold. Firstly, a bank entering a new market by M&A does not increase the number of banks in that market. Secondly, the reorganization processes following M&As are likely to affect the default rates, thus introducing a source of noise in our data.

Banks can enter local markets in two different ways: by branches or headquarters granting loans from outside the local market or by opening a branch. Both of these cases are well represented in our sample. Table I describes the numbers of entrants and incumbents and their relative proportion over the total sample. In Panel A entry is defined as acquiring a positive market share for the first time. By contrast, in Panel B we distinguish between entry with and without branches.[7]

[7] The large number of incumbents and entrants without branches reflects the lending activity of special credit institutions (SCIs). Over the sample period the regulation of the Italian banking system imposed a rather strict specialization of lending according to loan maturity. Most medium and long-term loans (i.e., with a maturity over 18 months) were granted by SCIs. These banks were not allowed to open branches and lending decisions were taken in their headquarters, although they used other banks' branches to collect loan applications, disburse funds and collect repayments. Short-term lending was mainly supplied by commercial banks, thrift institutions, cooperative and community banks. Both the specialization by maturity and the distinction between commercial banks and thrift institutions were cancelled by the new banking law passed in 1993.

Table I. Entry into Italian local credit markets between 1986 and 1991

Panel A shows the numbers of incumbents and entrants and their proportions of the total sample. We define a bank as an *Entrant* in a given market if its market share was equal to zero in 1986 and positive in 1991. Analogously a bank is an *Incumbent* in a given market if its market share was positive both in 1986 and in 1991. Panel B distinguishes between incumbents and entrants with and without branches. We define a bank as an *Entrant with Branches* in a given market if it did not have a branch in that market in 1986, but had at least one branch in 1991. An *Entrant with Loans* is a bank that had a market share equal to zero in 1986 and positive in 1991, but never opened a branch. A bank is an *Incumbent with Branches* in a given market if it had a branch in that market both in 1986 and in 1991; it is an *Incumbent with Loans* if it had a positive market share both in 1986 and 1991, but did not have a branch.

Panel A

	Proportion of total sample	1986		1991		Number of observations
		Loans	Branches	Loans	Branches	
Incumbent	78	yes	yes	yes	yes	1,852
		yes	no	yes	no	3,392
		yes	no	yes	yes	452
Entrant	22	no	no	yes	yes	41
		no	no	yes	no	1,538
Total	**100**					**7,275**

Panel B

	Proportion of total sample	1986		1991		Number of observations
		Loans	Branches	Loans	Branches	
Incumbent with Branches	25	yes	yes	yes	yes	1,852
Incumbent with Loans	47	yes	no	yes	no	3,392
Entrant with Branches	7	yes	no	yes	yes	452
		no	no	yes	yes	41
Entrant with Loan	21	no	no	yes	no	1,538
Total	**100**					**7,275**

We construct default rates, our dependent variable, as an average from 1993 to 1996 of the ratios between new bad loans at time T to loans not in default at time $T-1$. Loans are attributed to markets on the basis of customer location. The "adjusted" statistics on bad loans remove differences in banks' loan classification for multiple-bank borrowers. To cope with single-bank borrowers, we take the average ratio over several years. The initial period (1993) is the year of the business cycle trough, which was followed by a short recovery and a further slowdown. The final period (1996) marks the beginning of a steady recovery.

In computing the default rate, we used data on non-financial firms with bank debt ranging from 130,000 to 26,000,000 euros. Data below this range are rather noisy because borrowers with total loans and credit lines amounting to a value close to the reporting threshold may not be filed continuously by the CCR. Moreover, firms with bank debt greater than 26,000,000 euros are large and usually their loans are managed by banks' headquarters rather than by local branches.

Table II shows the definitions and sources of the variables we used in the different specifications of our empirical model. Table III displays their descriptive statistics.

3. Bad Loans, Entry and Information

3.1. THE BASIC ECONOMETRIC MODEL

We model the loan default rate of a bank in a given local market as a function of entry and information, plus two sets of dummy variables for banks' and markets' characteristics. We expect entrants to face more severe problems of adverse selection, since they will receive applications from all the firms that have previously been rejected by incumbents. Moreover, the adverse selection effect faced by incumbents should be inversely correlated with their knowledge of local markets. We estimate the following equation using weighted least squares logit regression for grouped data:[8]

$$\textit{Default Rate}_{ij,T_2} = \alpha_i B_i + \beta \textit{ Market Share}_{ij,T_0} + \gamma \textit{ Entry}_{ij} + \varphi_j P_j + \varepsilon_{ij} \qquad (1)$$

where *Deafult Rate*$_{ij,T_2}$ is the logistic transformation of the default rate of bank i's loans in market j at time T_2 (i.e., the average over the period 1993–1996). Local market characteristics are accounted for by the dummy P_j, while B_i is bank i's fixed effect. The variable *Entry*$_{ij}$ is a dummy that equals 1 when bank i enters market j in the period between T_0 and T_1 (1986 and 1991, respectively). *Entry*$_{ij}$ is equal to 1 if bank i's share in market j is 0 in 1986 and positive in 1991. Finally, we chose banks' loan market share in the initial period as a proxy for its knowledge of local market characteristics.[9]

[8] This estimation method is chosen because we are dealing with proportion data, i.e., the fraction of loans in a local market that defaults in a given time interval. The dependent variable is continuous and ranges from zero to one. Applying the logistic transformation to the dependent variable allows it to range over all real values (the logistic transformation of p is given by $\ln(p/1 - p)$). Since the variance of the default rate is inversely correlated with the size of total bank lending in the market under consideration, weighted least square estimation is needed in order to avoid heteroskedasticity problems. The default rate p was appropriately adjusted to compute its logistic transformation, which is not feasible if $p = 1$ or $p = 0$. Defining p as z/k, where z are the new bad loans and k the prior-year stock of loans not in default, the correction sets $z' = 0.001 * k$ whenever $z = 0$ and $k' = k + 0.001 * z$ whenever $k = z$.

[9] A large market share may also be associated with monopolistic power, which in turn may affect risk-taking. The exertion of market power, however, depends on the overall market structure, which

Table II. Description of the variables

The table displays the description and sources of the dependent and independent variables. The (D) notation indicates a dummy variable. CCR stands for the Italian Central Credit Register; SR BoI stands for Supervisory Reports to the Bank of Italy; ISTAT is the Italian National Institute for Statistics; IT stands for Istituto Tagliacarne, the research unit of the Italian Chambers of Commerce.

Symbol	**Description**	**Source**
	Dependent Variable	
Default Rate$_{ij}$	Bad loans at time T over loans not in default at $T-1$, of bank i in market j. Average from 1993 to 1996. In the case of multiple bank relationships, if not all the banks agree to classify a borrower as bad, the overall exposure of this borrower is reported in default if: (i) the borrower is reported as insolvent by a bank that accounts for at least 70 per cent of the exposure to the banking system; (ii) the borrower is reported as insolvent by two or more banks that account for at least 10 per cent of the total exposure to the banking system. The original default rate was corrected in order to compute its log-odds ratio. The correction was made as follows: denoting the default rate as $p = z/k$, where z is new bad loans and k the prior-year stock of loans not in default, we set $z' = 0.001 * k$ whenever $z = 0$ and $k' = k + 0.001 * z$ whenever $k = z$.	CCR
	Explanatory Variables	
	Bank – market variables	
Market Share	Market share of bank i in market j in 1986.	SR BoI
Pre Entry Market Share	Market share of bank i in market j the year before entry. Equals zero if the bank is an incumbent.	SR BoI
Entry (D)	Dummy variable equal to 1 if bank i passed from a zero to a positive market share in market j in the period 1986–1991.	SR BoI
Entry with Branches (D)	Dummy variable equal to 1 if bank i entered with at least one branch market j in the period 1986–1991.	
	SR BoI	
Entry with Loans (D)	Dummy variable equal to 1 if bank i entered, without opening a branch, market j in the period 1986–1991.	SR BoI
Incumbent with Loans (D)	Dummy variable equal to 1 if bank i was an incumbent without branches in market j in the period 1986–1991.	SR BoI
Distance	Distance between the entrant's target market and the bank's headquarter. Equals zero if the bank is incumbent.	
Deposits per Branch	Amount of deposits per branch in the entrant's target market. Equals zero if the bank is incumbent	SR BoI

Table II. Description of the variables (continued)

Symbol	Description	Source
	Market variables	
Turnover (D)	Equals 1 if customer turnover (i.e. the ratio between new customers in 1986 to existing customers in 1985) in market j exceeds the 75th percentile.	CCR
Market Size	Log of the population in market j in 1986.	ISTAT
Number of Banks	Log of the number of operating banks in market j in 1986.	SR BoI
Herfindahl	Herfindahl index of market j computed on loans, based on the location of the borrower, in 1986.	SR BoI
Interest Rate	Average interest rate in market j in 1986.	SR BoI
Court Efficiency	Log of the average number of days needed to complete bankruptcy proceedings in market j during the period 1983–1986.	ISTAT
1986 GDP	Per capita value added of market j in 1986.	IT
1993 GDP Growth Rate	Rate of growth of value added in real terms in market j in 1993.	IT
Number of Entrants	Log of one plus the number of entrants in market j during the period 1986–1991.	SR BoI
	Bank variables	
Bank Size	Log of total assets of bank i in 1986.	SR BoI
Capital	Ratio of capital to total assets of bank i in 1986.	SR BoI
Bad Loans	Ratio of the stock of bad loans to the stock of loans not in default of bank i in 1986.	SR BoI
Profits	Pretax return on equity of bank i in 1986.	SR BoI
Cooperative (D)	Dummy variable equal to 1 if bank i is a cooperative bank.	SR BoI
Community (D)	Dummy variable equal to 1 if bank i is a community bank.	SR BoI
Special Credit Inst.(D)	Dummy variable equal to 1 if bank i is a special credit institution.	SR BoI
Thrift (D)	Dummy variable equal to 1 if bank i is a thrift institution.	SR BoI

Table III. Descriptive statistics

The table displays the descriptive statistics for the dependent and independent variables used in the estimations described in Tables IV and VII. The database has 7,275 observations, referring to 729 banks and 95 local markets. The (D) notation indicates a dummy variable.

Symbol	**Mean**	**Min.**	**Max.**	**St. dev.**
	Dependent Variable			
Default Rate	0.048	0.001	0.999	0.106
	Explanatory Variables			
Bank – market variables				
Market Share	0.012	0.000	0.461	0.034
Pre-Entry Market Share	0.001	0.000	0.223	0.005
Entry (D)	0.217	0.000	1.000	0.412
Entry with Branches (D)	0.068	0.000	1.000	0.251
Entry with Loans (D)	0.211	0.000	1.000	0.408
Incumbent with Loans (D)	0.466	0.000	1.000	0.499
Market variables				
Turnover (D)	0.292	0.000	1.000	0.455
Market Size	13.228	11.433	15.191	0.829
Number of Banks	4.851	3.296	5.979	0.546
Herfindahl	0.083	0.029	0.261	0.042
Interest Rate	14.823	12.660	17.960	1.311
Court Efficiency	7.652	7.389	8.125	0.177
1986 GDP	3.313	1.369	5.122	0.527
1993 GDP Growth Rate	−1.580	−4.920	1.737	1.401
Bank variables				
Bank Size	7.798	1.692	11.275	2.050
Capital	0.059	0.005	0.204	0.033
Bad Loans	0.063	0.000	0.472	0.034
Profits	0.262	0.012	2.979	0.156
Cooperative (D)	0.197	0.000	1.000	0.397
Community (D)	0.126	0.000	1.000	0.332
Special Credit Inst. (D)	0.195	0.000	1.000	0.397
Thrift (D)	0.204	0.000	1.000	0.403

Our empirical analysis spans a decade, which is consistent with earlier studies. For instance Shaffer (1998) finds that from year 3 to year 9 of their existence *de*

we control for through the dummies P_j. Theoretical arguments supporting the choice of market share as a proxy for information can be drawn from the results obtained by Sharpe (1990), Dell'Ariccia et al. (1999) and Marquez (2002).

novo banks have significantly higher net chargeoff ratios than mature banks, and argues that this is due to adverse selection. Using a long period between entry and the measurement of default rates can be justified on the ground that it is very difficult to establish when a bank actually starts its operations on an economically meaningful scale in a newly entered market. In addition, borrowers' defaults usually occur after several months, often a few years, from the time the loans were granted, especially when there are no macroeconomic shocks. Over the period 1987–1993 the Italian economy enjoyed sustained growth, which reduced the probability of default even of high-risk borrowers. To test the robustness of the results, we also run the regressions using the 1993 default rate as the dependent variable without finding any substantial change in the results.

The above definition of entry is very broad: the acquisition of a positive market share may be episodic and need not reflect a strategic entry decision. An alternative is to define entry as the opening of a branch by an outside bank. When a bank enters a market by opening a branch, it presumably has more information about local market conditions and potential new borrowers than when it enters by granting loans from outside. This difference in knowledge may be due to the bank already having some customers in that market or to preliminary market research.[10] Opening a new branch involves incurring fixed costs that must be justified by the expectation of reaching a critical mass of loans, and these expectations must be supported by information. Moreover, when a bank opens a new branch it is likely to provide payment services to its borrowers. Black (1975) and Fama (1985) argue that this may greatly help banks in their monitoring activity. Mester et al. (2005) provide empirical evidence that transaction account information actually improves monitoring. We therefore compare the effects on the loan default rate of different types of entry characterized by different levels of information. Equation (1) is modified as follows:

$$\begin{aligned} \textit{Default Rate}_{ij,T_2} = {} & \alpha_i B_i + \beta \textit{ Market Share}_{ij,T_0} + \phi \textit{ Entry with Branches}_{ij} + \\ & + \nu \textit{ Entry with Loans}_{ij} + \rho \textit{ Incumbent with Loans}_{ij} + \\ & + \delta_j P_j + \varepsilon_{ij} \end{aligned} \tag{2}$$

where *Entry with Branches*$_{ij}$ indicates whether bank i opened a branch in market j, *Entry with Loans*$_{ij}$ indicates that bank i entered market j and acquired at least one new customer but without opening a branch; and *Incumbent with Loans*$_{ij}$ is a dummy that is equal to 1 if bank i is an incumbent in market j but never opened a branch. We expect all three dummies to have positive coefficients. In particular, the effect on the default rates should be larger for banks that entered a market by lending from outside than for those that opened a branch.

[10] In our sample 452 out of 493 episodes of entry with branches refer to banks that were already lending in the markets in question.

3.2. RESULTS

We estimate Equations (1) and (2) using weighted least squares logit regression for grouped data. Estimation results are shown in the first two columns of Table IV.

The entry dummies are positive and highly significant under both the definitions adopted. The coefficient associated with market share is negative and significantly different from zero: the greater the information about the economic conditions in the local market and the pool of customers, the lower the default rate. The positive coefficient associated with *Incumbent with Loans* implies that granting loans without being on site leads to higher default rates.

As well known, in a logit regression for grouped data it is not possible to interpret the regression coefficients as the partial derivatives of the conditional expectation of the default rate with respect to the associated independent variables. In order to assess the economic significance of the coefficients, we estimate the predicted loan default rates for different values of the relevant independent variables using Duan's (1983) smearing estimator. Using the estimated regression coefficients, we compute the regression residuals as the difference between the dependent variable and the linear predictions (Manning, 1998). Then we compute the predicted values for each observation as the sum of the residuals and the product of the vector of the regression coefficients and the vector of the covariates evaluated at the mean value, except for the variables of interest. Finally, we apply the inverse of the logistic transformation and take the average.

Table V reports the predicted default rates based on the regression coefficients of Equations (1) and (2) (Panels A and B, respectively). We distinguish between incumbents with a small market share (the bottom decile of the distribution) and incumbents with a large market share (the top decile). Panel A shows that loans granted by new entrants (defined here as banks that previously did not lend in the province) end up in default with a 7.44 per cent probability, compared with 3.08 per cent for the loans granted by incumbents with a large market share (defined as the 90th percentile of the distribution of incumbents). Incumbents with a small market share (defined as the 10th percentile) experience an intermediate default rate (3.59 per cent) that is below the sample mean. The differences in the predicted probabilities between the two types of incumbents can be explained by the different endowments of information. Entrants combine scarce information about local market conditions with greater exposure to adverse selection.

The results reported in Panel B distinguish between banks that enter a new market by opening a branch and those that enter by granting loans. The different categories of banks are ranked in ascending order with respect to the predicted default rate of their loan portfolio. The ranking would be exactly the same if they were ordered according to the amount of information they presumably have about local market conditions or their ability to monitor borrowers. Incumbent banks with a large market share and with branches are best placed to assess the creditworthiness of their borrowers and to monitor them. Their predicted default rate, equal to

Table IV. Determinants of the loan default rate: The fixed effect model

The table reports regression coefficients and associated standard errors, robust to heteroskedasticity, for the fixed effect model. The regressions are estimated with weighted least square logit for grouped data; all the regressions are run on a dataset with 7,275 observations, referred to 729 banks and 95 local markets. The dependent variable is the logistic transformation of the loan default rate. The independent variables are described in Table II. The (D) notation indicates a dummy variable. Dummy variables for banks and markets are not reported. The panel below shows the Wald tests for the differences between the relevant coefficients.

Equation	(1)	(2)	(3)	(4)
Constant	−3.057***	−3.097***	−3.085***	−3.115***
	0.049	*0.046*	*0.046*	*0.046*
Market Share	−3.421***	−2.377***	−2.374***	−2.159***
	0.120	*0.125*	*0.125*	*0.126*
Entry (D)	0.964***			
	0.081			
Entry with Branches (D)		0.187***	0.238***	0.289***
		0.032	*0.040*	*0.036*
Entry with Branches			−0.127**	
∗ Turnover (D)			*0.061*	
Entry with Branches				−3.820***
∗ Pre-Entry Market Share				*0.760*
Entry with Loans (D)		1.392***	1.589***	1.392***
		0.083	*0.101*	*0.082*
Entry with Loans			−0.551***	
∗ Turnover (D)			*0.162*	
Incumbent with Loans (D)		0.687***	0.732***	0.775***
		0.027	*0.032*	*0.029*
Incumbent with Loans			−0.100**	
∗ Turnover (D)			*0.042*	
Incumbent with Loans				−5.331***
∗ Market Share				*0.573*
Adj. R-squared	*0.623*	*0.659*	*0.659*	*0.664*

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Equation	Test	F-statistic	P-value
(2)	*Entry with Branches – Entry with Loans = 0*	*F(1, 6,448) = 196.73*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 6,448) = 175.05*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 6,448) = 71.65*	*0.000*
(3)	*Entry with Branches – Entry with Loans = 0*	*F(1, 6,446) = 160.98*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 6,446) = 134.05*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 6,446) = 55.08*	*0.000*
(4)	*Entry with Branches – Entry with Loans = 0*	*F(1, 6,445) = 162.69*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 6,445) = 111.56*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 6,445) = 70.71*	*0.000*

Table V. Economic significance: The fixed effect model

The table reports the smearing estimates (Duan, 1983) of the effect of entry on the loan default rate. We obtain the predicted default rates shown in Panel A as follows. Using the coefficients reported in the first column of Table IV, we compute the regression residuals as the difference between the dependent variable and the linear predictions (Manning, 1998). Then we compute the predicted values for each observation by adding to the residuals the product of the vector of the regression coefficients times the vector of the covariates evaluated at the mean value, except for the variables of interest. That is, we set *Market Share* = 90th percentile of the distribution of *Market Share* for the incumbents and *Entry* = 0, if computing the default rate of the incumbents with a large market share; *Market Share* = 10th percentile and *Entry* = 0, if computing the default rate of the incumbents with a small market share; and *Market Share* = 0 and *Entry* = 1, if computing the default rate of entrants. Then we apply the inverse of the log-odds transformation and lastly we take the average, obtaining the predicted default rate. The results shown in Panel B are obtained using the regression coefficients reported in the third column of Table IV.

Panel A	
	Predicted default rate
Incumbents with large market share	3.08
Incumbents with small market share	3.59
Entrants	7.44
Panel B	
	Predicted default rate
Incumbents with large market share and with branches	2.00
Incumbents with small market share and with branches	2.58
Entrants with branches	2.90
Incumbents with large market share and without branches	4.40
Incumbents with small market share and without branches	4.53
Entrants without branches	7.78

2 per cent, is far below the sample mean. This informational advantage declines if the market share is small: entrants with branches experience a default rate 0.9 percentage points higher than incumbents with branches and a large market share. Having a branch on site seems to be very important in order to avoid credit losses. According to our estimates, even incumbent banks with a relatively large market share, but without a branch on site, are likely to have high default rates (4.40 per cent) that is more than twice that of incumbents with a large market share and branches. The winner's curse effect is likely to be stronger for lenders that grant loans from outside. This, plus the scarce information about local market conditions,

explains the extremely high default rate (7.78 per cent) on the loan portfolio of banks that enter without opening a branch.

3.3. EXTENSIONS

Our basic empirical specification can be easily extended to test other implications suggested by theory and to check for the robustness of our results.

3.3.1. *Customer Turnover*

Dell'Ariccia et al. (1999) argue that the comparative disadvantage of less informed banks should be mitigated by a high turnover of bank customers. In each period the pool of potential borrowers is composed of those previously rejected and of first-time applicants. The larger the latter component, the lower entrants' informational disadvantage. We test this hypothesis by introducing in Equation (2) an interaction variable between the entry dummy and a dummy identifying high turnover markets. For each market we define the customer turnover as the ratio between new customers in 1986 and existing customers in 1985. We set the dummy variable *Turnover* equal to 1 when the customer turnover exceeds the 75th percentile and zero otherwise, and interact it with the entry variables. We expect a negative coefficient for the interaction and a stronger effect of the entry dummy on the default rate than in the basic specification. Column (3) of Table IV reports the results, which are again consistent with the predictions. The effect of customer turnover is also economically significant: computing the predicted default rates, we find that entrants without branches in high-turnover markets have a loan default rate of 5.92, compared with 8.95 for entrants in low-turnover markets.

3.3.2. *Pre-entry Market Share*

The difference between entry with and without branches suggests a further possible test. The informational disadvantage of entrants with branches should be lower if they were already granting loans in the market they entered. This reduction of the informational disadvantage should decrease with their market share (our proxy for information) before entry. By the same token the default rate of incumbents without branches should decrease with their market share. We test this hypothesis by interacting *Entry with Branches* with the pre-entry market share (*Pre-Entry Market Share*) and *Incumbent with Loans* with the initial period market share. The interactions should reduce the coefficient of *Market Share* and increase that of *Entry with Branches* and *Incumbent with Loans* compared with Equation (2). At the same time the two interactions are expected to have a negative effect on the default rate.

Estimation results are shown in the last column of Table IV. The coefficients associated with the two interaction terms are significant and have the expected sign. In order to highlight the role of information in reducing the loan default rate, we compute the predicted values using the regression coefficients. Entrants with branches, but with a small pre-entry market share, experience loan loss rates 6 per cent higher than those with a large pre-entry market share. Within the group of incumbents with branches, there is a large difference between those with large market share and those with a small one (10 per cent).

4. Robustness

The results of our estimates support the hypothesis that entrants have a smaller amount of information regarding the pool of potential borrowers and therefore have a higher probability of granting credit to bad applicants. In our model, being an entrant can also be interpreted as being inexperienced and the coefficients associated with the variables *Entry*, *Entry with Branches* and *Entry with Loans* can be interpreted as the price a bank has to pay for its scarce knowledge of its potential pool of borrowers. The same applies to the interpretation of the coefficient of *Incumbent with Loans*, as far as having a branch on site implies a grater ability to collect information.

On the other hand if banks are forward-looking and the entry decision is taken on the basis of the expected default rates, our key variables are potentially endogenous. Insofar the bias is determined by the fact that banks that anticipate excessively high default rates do not enter new markets, it is likely that the coefficients of the variables *Entry* and *Entry with Loans* underestimate the effect of asymmetric information. The same could be true for the variable *Entry with Branches*, unless the decision to open a branch is taken precisely to avoid high default rates. In this case the bias can go in the opposite direction. In order to control for the endogeneity problem we estimated Equations (1) and (2) with instrumental variables. In Equation (1) the potentially endogenous variable *Entry* is instrumented using the distance between the entrant's target market and the bank's headquarter (*Distance*). In Equation (2) the two potentially endogenous variables, *Entry with Loans* and *Entry with Branches*, are instrumented, respectively, with *Distance* and with the amount of deposits per branch in the entrant's target market (*Deposits per Branch*). Our underlying assumption is that these two variables are correlated with the entry decision,[11] but uncorrelated with the expected default rate. The results of the instrumental variables estimations are reported in Table VI. The coefficients of the potential endogenous variables are not substantially different from those obtained in the previous regressions.

5. Bank and Market Characteristics

As an extension of the basic model we replace bank and market fixed effects with appropriate controls in order to have insights on bank and market characteristics

[11] The results of the first stage regressions, not reported for brevity, but available from the authors upon request, support this hypothesis.

Table VI. Determinants of the loan default rate: The fixed effect model with Instrumental Variables

The table reports regression coefficients and associated standard errors, robust to heteroskedasticity, for the IV fixed effect model. The regressions are estimated with IV weighted least square logit for grouped data; all the regressions are run on a dataset with 7,275 observations, referred to 729 banks and 95 local markets. The dependent variable is the logistic transformation of the loan default rate. The independent variables are described in Table II. In (1) the potentially endogenous variable *Entry* is instrumented with the distance between the entrant's target market and the bank's headquarter (*Distance*). In (2) the two potentially endogenous variables, *Entry with Branches* and *Entry with Loans*, are instrumented, respectively, with the amount of deposits per branch in the entrant's target market (*Deposits per Branch*) and the distance between the entrant's target market and the bank's headquarter. The (D) notation indicates a dummy variable. Dummy variables for banks and markets are not reported. The panel below shows the Wald tests for the differences between the relevant coefficients.

Equation	(1)	(2)
Constant	−3.057***	−3.095***
	0.049	*0.046*
Market Share	−3.405***	−2.392***
	0.120	*0.128*
Entry (D)	1.070***	
	0.133	
Entry with Branches (D)		0.162***
		0.035
Entry with Loans (D)		1.475***
		0.138
Incumbent with Loans (D)		0.687***
		0.027
Adj. R-squared	*0.623*	*0.658*

*** , ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Equation	Test	F-statistic	P-value
(2)	*Entry with Branches - Entry with Loans = 0*	*F(1, 6,448) = 92.61*	*0.000*
	Entry with Branches - Incumbent with Loans = 0	*F(1, 6,448) = 178.10*	*0.000*
	Entry with Loans - Incumbent with Loans = 0	*F(1, 6,448) = 34.74*	*0.000*

that affect the loan default rate. This procedure also allows us to test two additional hypotheses: the first is that if credit screening is imperfectly correlated across banks, the average quality of loans should decrease as the number of lenders in the market increases; the second is that borrower turnover, though mitigating the relative disadvantage of entrants, should be positively correlated with the average market default rate as it increases the share of borrowers known only through creditworthiness tests.

The empirical models have the same structure as those previously described. Their specification can be found in Table VII. We replace the bank an market fixed effects with two sets of covariates accounting for the initial conditions at time T_0. The two hypotheses are tested by introducing the dummy *Turnover* and the variable *Number of Banks*, computed as the natural log of the number of banks operating in each market at the beginning of the sample period. The coefficients of both these variables are expected to have a positive sign.

Market controls include size (*Market Size*), measured by the natural log of the population and per capita GDP (*1986 GDP*), both taken at the initial period values. Since an economic downturn might have differently affected default rates across markets, we also introduce the 1993 province-level GDP growth rates (*1993 GDP Growth Rate*).

To control for the structure of the local credit market, the explanatory variables also include the Herfindahl-Hirschman concentration indexes (*Herfindahl*), computed on loan market shares in the provinces. There are several reasons for introducing this variable. First, competition, may affect the incentives to gather information about borrowers, either by reducing relationship lending (Chan et al., 1986; Petersen and Rajan, 1995), or by increasing it (Boot and Thakor, 2000).[12] Second, high market concentration is frequently associated with high loan interest rates (Berger and Hannan, 1989), therefore attracting high risk borrowers (Stiglitz and Weiss, 1981). Consequently, market concentration should be positively correlated with default rates (Boyd and De Nicolò, 2005). A further reason to control for the level of competition is the use of banks' market shares as proxies for information.

We account for the average quality of borrowers in the market by introducing the variable *Interest Rate*, defined as the average market interest rate. In Italy court proceedings take a long time to produce an outcome (Bianco et al., 2005). Differences in court efficiency in enforcing bankruptcy procedures may be reflected in opportunistic behavior on the part of borrowers (Shleifer and Vishny, 1993). Therefore, we include the variable *Court Efficiency*, the average number of days needed to complete bankruptcy proceedings; the expected sign of its coefficient is positive.

Regarding banks' characteristics, we control for size, efficiency and leverage. Bad management may have poor skills in credit scoring, be unable to appraise the value of collateral and have difficulty in monitoring. Banks that had a high proportion of bad loans in the initial period (*Bad Loans*) are also likely to have a high default rate in later periods. Overall efficiency is proxied by gross returns on equity (*Profits*) and is expected to have a negative coefficient. The loan default

[12] Degryse and Ongena (2004) empirically investigates the effect of competition on relationship lending, finding that they are not necessarily inimical. Elsas (2005) finds a non-monotonical relationship between market concentration and relationship lending.

Table VII. Determinants of the loan default rate: The covariate model

The table reports regression coefficients and associated standard errors, robust to heteroskedasticity, for the covariate model. The regressions are estimated with weighted least square logit for grouped data; all the regressions are run on a dataset with 7,275 observations, referred to 729 banks and 95 local markets. The dependent variable is the logistic transformation of the loan default rate. The independent variables are described in Table II. The (D) notation indicates a dummy variable. The panel below shows the Wald tests for the differences between the relevant coefficients.

Equation	(1)	(2)	(3)	(4)
Constant	−8.325***	−9.111***	−9.132***	−8.687***
	0.582	*0.564*	*0.565*	*0.560*
Market Share	−2.063***	−1.170***	−1.164***	−0.845***
	0.130	*0.133*	*0.133*	*0.135*
Entry (D)	0.902***			
	0.100			
Entry with Branches (D)		0.263***	0.281***	0.335***
		0.038	*0.048*	*0.043*
Entry with Branches ∗ Turnover (D)			−0.042	
			0.073	
Entry with Branches ∗ Pre-Entry Market Share				−1.999***
				0.778
Entry with Loans (D)		1.363***	0.281***	1.365***
		0.104	*0.048*	*0.103*
Entry with Loans ∗ Turnover (D)			−0.506**	
			0.210	
Incumbent with Loans (D)		0.719***	0.733***	0.862***
		0.033	*0.039*	*0.035*
Incumbent with Loans ∗ Turnover (D)			−0.031	
			0.052	
Incumbent with Loans ∗ Market Share				−6.928***
				0.618
Turnover (D)	0.399***	0.417***	0.429***	0.401***
	0.023	*0.023*	*0.024*	*0.022*
Market Size	−0.021	0.023	0.021	0.022
	0.027	*0.026*	*0.026*	*0.026*
Number of Banks	0.328***	0.301***	0.303***	0.280***
	0.055	*0.054*	*0.054*	*0.053*
Herfindahl	6.380***	6.312***	6.312***	6.023***
	0.315	*0.305*	*0.315*	*0.304*
Interest Rate	0.169***	0.164***	0.165***	0.158***
	0.011	*0.011*	*0.011*	*0.011*

Table VII. Determinants of the loan default rate: The covariate model (continued)

Equation	(1)	(2)	(3)	(4)
Court Efficiency	0.323***	0.333***	0.336***	0.307***
	0.057	*0.055*	*0.055*	*0.055*
1986 GDP	−0.248***	−0.192***	−0.194***	−0.194***
	0.039	*0.038*	*0.038*	*0.038*
1993 GDP Growth Rate	−0.052***	−0.055***	−0.055***	−0.051***
	0.007	*0.007*	*0.007*	*0.007*
Bank Size	−0.048***	−0.050***	−0.051***	−0.049***
	0.008	*0.008*	*0.008*	*0.008*
Capital	−4.908***	−4.481***	−4.473***	−4.720***
	0.414	*0.401*	*0.402*	*0.402*
Bad Loans	4.611***	4.648***	4.628***	4.605***
	0.397	*0.385*	*0.386*	*0.382*
Profits	−1.332***	−1.290***	−1.285***	−1.283***
	0.069	*0.067*	*0.067*	*0.066*
Special Credit Institutions (D)	0.264**	−0.126***	−0.059	−0.129
	0.032	*0.036*	*0.037*	*0.036*
Thrift	0.069**	−0.011	−0.031	−0.011***
	0.028	*0.028*	*0.028*	*0.028*
Cooperative (D)	−0.069**	−0.091***	−0.091***	−0.092***
	0.033	*0.032*	*0.032*	*0.032*
Community (D)	−0.348***	−0.291***	−0.291***	−0.268***
	0.085	*0.083*	*0.083*	*0.082*
Adj. R-squared	*0.341*	*0.383*	*0.384*	*0.394*

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Equation	Test	F-statistic	P-value
(2)	*Entry with Branches – Entry with Loans = 0*	*F(1, 7,254) = 196.58*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 7,254) = 100.19*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 7,254) = 37.91*	*0.000*
(3)	*Entry with Branches – Entry with Loans = 0*	*F(1, 7,252) = 88.91*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 7,252) = 105.66*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 7,252) = 23.16*	*0.000*
(4)	*Entry with Branches – Entry with Loans = 0*	*F(1, 7,251) = 86.72*	*0.000*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 7,251) = 62.74*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 7,251) = 38.69*	*0.000*

rate may also be affected by moral hazard. Banks with low returns may be tempted to gamble, assuming too much risk in order to remain in the market. This temptation increases with leverage, owing to the effect of deposit insurance and limited liability (Brander and Lewis, 1986; Dewatripont and Tirole, 1994). Therefore, the variable *Capital*, defined as the ratio of equity capital to total assets, is expected to have a negative coefficient.

We also include a set of dummy variables indicating the institutional status of the bank (i.e., thrift institution, special credit institution, cooperative bank or community bank).

The estimation results for the covariate model are reported in Table VII and organized in the same way as Table IV. Replacing market and bank fixed effects with the covariates does not alter the sign or significance of the entry dummies and their interactions with market share, pre-entry market share and customer turnover. Again, the overall pattern of coefficients is consistent across the different specifications. The data support our hypotheses regarding the effects of the number of banks operating in a market and the existence of a high customer turnover: both *Number of Banks* and *Turnover* have a positive and significant coefficient. Banks operating in markets characterized by a large number of lenders experience higher default rates, as predicted by theory, and high customer turnover increases loan losses. Instead, *Herfindahl* has a positive sign, which is consistent with the prediction of Boyd and De Nicolò (2005).[13]

All the remaining coefficients on market covariates have the expected signs. The hypotheses regarding bank characteristics are also confirmed. Banks with high leverage, less capable managers and low efficiency experience significantly higher default rates, as confirmed by the signs associated with the variables *Capital*, *Bad Loans* and *Profits*.

Finally, the estimates are consistent with the view that the strong local roots of cooperative and community banks generate information that gives them an advantage in screening and monitoring activities. The finding that these intermediaries experience below-average loan default rates is consistent with other empirical studies performed using different samples and different methodologies (Cannari and Signorini, 1997).

6. Loan Default Rates and Banks' Organizational Structure

In the previous sections we provided extensive evidence of a positive relationship between entry and loan default rates. Our empirical specification, together with our robustness checks, indicates that entrants experience higher default rates because of their informational disadvantage with respect to incumbents. Our findings may

[13] Since we are aware of the potential problems due to multicollinearity with other variables we have estimated all the equations without the variable *Herfindahl*, with no qualitative change and a small increase of the standard errors of some coefficients.

nevertheless be consistent with an alternative interpretation based on organizational differences across banks.

Stein (2002) argues that a bank's organizational structure can play an important role in defining the procedures used to assess the creditworthiness of potential customers. Small banks, in which just a few managers are responsible for the evaluation of the quality of credit applicants, provide high incentives to collect "soft" information (i.e., information that can not be directly verified by anyone else but the agent who produces it). Conversely, large hierarchical banks base their lending decisions mainly on "hard" information that can be easily transmitted and verified along the hierarchical structure. It follows that large banks may have difficulty in assessing the creditworthiness of small businesses, for which "soft" information is likely to be particularly important.

In our basic specification we introduced bank fixed effects, so that, in principle, the relationship between loan default rates and entry should not be attributed to the organizational structure of the banks involved. Nevertheless, small banks are over-represented in our sample and the fixed effects could be inadequate to control fully for differences in organizational structures. Out of the 729 banks in the sample, 486 are small community banks, typically operating in one local market, and 152 are cooperative or thrift institutions. At the same time, only 37 out of 493 cases of entry with branches involve community banks.

We check for the possible bias induced by the composition of our sample by restricting the analysis to those banks that where incumbents (with branches) in at least five provinces in 1986 and, at the same time, entered with branches in at least one local market. The reduced sample has 2,641 observations and refers to 42 banks, none of which is a community bank. In this sample, the standard deviation of the log of total assets is equal to 1.05, compared with 2.05 for the full sample. We exclude special credit institutions entirely. This implies giving consideration only to short-term loans, a more homogeneous product, and reducing the proportion of incumbents without branches.

Table VIII reports the results of the estimation using this reduced sample. The signs and the significance of the coefficients are not altered, although the estimated coefficients of *Entry with Loans* and *Incumbent with Loans* are not statistically different. Table IX shows the predicted default rates.

7. Consequences of Entry on Incumbents' Loan Default Rates

A natural extension of the analysis is to assess the consequences of entry on incumbents' default rates. One can think of two possible effects, working in opposite directions. The first is related to incumbents' customers. Lacking information, entrants offer an interest rate that reflects the average credit quality of borrowers in the market. They consequently attract incumbents' riskier customers, who are charged high interest rates. This would imply a negative relationship between the number of entrants in a given market and incumbents' default rate. The second

Table VIII. Estimates with the reduced sample

The table reports regression coefficients and associated standard errors, robust to heteroskedasticity, for the fixed effect model. The regressions are estimated with weighted least square logit for grouped data; all the regressions are run on a dataset with 2,641 observations, referred to 42 banks and 95 local markets. All the banks in the sample were incumbents with branches in at least five provinces in 1986, and entered with branches in at least one local market. The dependent variable is the logistic transformation of the loan default rate. The independent variables are described in Table II. The (D) notation indicates a dummy variable. Dummy variables for banks and markets are not reported. The panel below shows the Wald tests for the differences between the relevant coefficients.

Equation	(1)	(2)	(3)	(4)
Constant	−3.206***	−3.241***	−3.235***	−3.247***
	0.061	*0.059*	*0.059*	*0.059*
Market Share	−3.054***	−2.226***	−2.237***	−2.219***
	0.157	*0.164*	*0.164*	*0.163*
Entry (D)	0.626***			
	0.203			
Entry with Branches (D)		0.228***	0.279***	0.293***
		0.048	*0.060*	*0.057*
Entry with Branches			−0.137	
∗ Turnover (D)			*0.095*	
Entry with Branches				−4.622**
∗ Pre-Entry Market Share				*2.183*
Entry with Loans (D)		0.997***	1.076***	1.001***
		0.224	*0.286*	*0.223*
Entry with Loans			−0.210	
∗ Turnover (D)			*0.457*	
Incumbent with Loans (D)		0.772***	0.803***	0.892***
		0.054	*0.065*	*0.065*
Incumbent with Loans			−0.088	
∗ Turnover (D)			*0.106*	
Incumbent with Loans				−14.433***
∗ Market Share				*4.491*
Adj. R-squared	*0.702*	*0.725*	*0.725*	*0.727*

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Equation	Test	F-statistic	P-value
(2)	*Entry with Branches – Entry with Loans = 0*	*F(1, 2,501) = 11.50*	*0.001*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 2,501) = 67.50*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 2,501) = 0.98*	*0.326*
(3)	*Entry with Branches – Entry with Loans = 0*	*F(1, 2,499) = 9.63*	*0.002*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 2,499) = 54.24*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 2,499) = 0.23*	*0.634*
(4)	*Entry with Branches – Entry with Loans = 0*	*F(1, 2,498) = 7.55*	*0.006*
	Entry with Branches – Incumbent with Loans = 0	*F(1, 2,498) = 39.85*	*0.000*
	Entry with Loans – Incumbent with Loans = 0	*F(1, 7,251) = 0.88*	*0.347*

Table IX. Economic significance with the reduced sample

The table reports the smearing estimates (Duan, 1983) of the effect of entry on the loan default rate. We obtain the predicted default rates shown in Panel A as follows. Using the coefficients reported in the first column of Table VII, we compute the regression residuals as the difference between the dependent variable and the linear predictions (Manning, 1998). Then we compute the predicted values for each observation by adding to the residuals the product of the vector of the regression coefficients times the vector of the covariates evaluated at the mean value, except for the variables of interest. That is, we set *Market Share* = 90th percentile of the distribution of *Market Share* for the incumbents and Entry = 0, if computing the default rate of the incumbents with a large market share; *Market Share* = 10th percentile and Entry = 0, if computing the default rate of the incumbents with a small market share; *Market share* = 0 and *Entry* = 1, if computing the default rate of entrants. Then we apply the inverse of the log-odds transformation and finally we take the average, obtaining the predicted default rate. The results shown in Panel B are obtained using the regression coefficients reported in the third column of Table VII.

Panel A	
	Predicted default rate
Incumbents with a large market share	3.31
Incumbents with a small market share	4.12
Entrants	6.78
Panel B	
	Predicted default rate
Incumbents with a large market share and with branches	2.22
Incumbents with a small market share and with branches	2.95
Entrants with branches	3.62
Incumbents with a large market share and without branches	5.68
Incumbents with a small market share and without branches	5.81
Entrants without branches	6.98

effect is related to first-time loan applicants. By increasing the number of banks, entry increases the probability that a bad borrower will be found creditworthy by at least one of them. Therefore, a large number of entrants should be correlated with higher default rates for incumbents.[14] In order to test which of these two effects prevails, we estimate the following equation:

$$\begin{aligned} \textit{Default Rate}_{ij,T_2} \;=\; & \alpha_i B_i + \beta\, \textit{Market Share}_{ij,T_0} + \eta\, \textit{Incumbents with Loans}_{ij} + \\ & + \phi\, \textit{Number of Entrants}_j + \gamma\, \textit{Turnover}_j + \\ & + \delta\, \textit{Number of Entrants}_j * \textit{Turnover}_j + \\ & + \textit{Market Controls} + \varepsilon_{ij} \qquad (3) \end{aligned}$$

[14] This is consistent with Broecker (1990) and Shaffer (1998). Gehrig (1998) points out that this may not be true if incumbents adjust their own screening to the new environment.

Table X. Effect of entry on incumbents' default rate

The table reports regression coefficients and associated standard errors, robust to heteroskedasticity, for the incumbents' default rate model. The regressions are estimated with weighted least square logit for grouped data; all the regressions are run on a dataset with 5,244 observations, referred to 726 banks and 95 local markets. All the banks in the sample were incumbents. The dependent variable is the logistic transformation of the loan default rate. The independent variables are described in Table II. The (D) notation indicates a dummy variable. Dummy variables for banks are not reported.

Equation	(1)	(2)
Constant	−8.458***	−8.043***
	0.620	*0.654*
Market Share	−2.442***	−2.420***
	0.152	*0.153*
Incumbents with Loans	0.674***	0.675***
	0.033	*0.033*
Number of Entrants	−0.191***	−0.224***
	0.036	*0.040*
Number of Entrants $*$ Turnover		0.100*
		0.051
Turnover (D)	0.337***	0.004**
	0.025	*0.170*
Market Size	0.071**	0.038
	0.030	*0.034*
Number of Banks	0.199***	0.232***
	0.056	*0.059*
Herfindahl	4.785***	4.733***
	0.351	*0.352*
Interest Rate	0.107***	0.108***
	0.011	*0.011*
Court Efficiency	0.350***	0.357***
	0.060	*0.060*
1986 GDP	−0.111**	−0.140***
	0.044	*0.046*
1993 GDP Growth Rate	−0.016**	−0.014*
	0.008	*0.008*
Adj. R-squared	*0.586*	*0.586*

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

where *Default Rate*$_{ij,T_2}$ is the logistic transformation of the default rate of incumbent bank i's loans in market j at time T_2; *Number of Entrants*$_j$ is the number of entrants in market j in the period between T_0 and T_1 and *Market Controls* stands for the set of market variables described in the previous section. Depending on which of the two effects prevails the variable *Number of Entrants* will have a positive or negative coefficient. Independently of the prevailing effect, incumbents in markets characterized by a high level of turnover should exhibit higher default rates, since in every period there is a larger proportion of borrowers whose risk level can be assessed only through the creditworthiness test. Actually, the second effect only works on new customers, so that introducing an interaction between the number of entrants and turnover allows us to disentangle the two effects.

Table X reports the results of the analysis of the effects of entry on incumbents' default rate. The two columns show the regression coefficients of the estimate of Equation (3), with and without the interaction between the number of entrants and turnover. The coefficient of the variable *Number of Entrants* is negative. The improvement in the quality of incumbents' loan portfolios – prompted by the switching of their riskier customers to the entrants – more than compensates for the worsening due to the higher probability of their assessing a high-risk new customer as creditworthy. The second column of the table shows, however, that the latter effect does exist and is significant. The coefficient of the interaction between the number of entrants and customer turnover is positive, while the absolute value of the coefficient of *Number of Entrants* increases. This is consistent with the hypothesis that, as far as first-time loan applicants are concerned, an increase in the number of banks in the market causes incumbents' default rate to rise. Finally, this further specification can also be interpreted as a robustness check of the result that more information reduces the risks taken by banks. Both the coefficients of *Incumbent with Loans* and of *Market Share* are significant and with the same sign as in the previous specifications.

8. Conclusions

This paper has explored the links between entry and the default rates of loans granted by entrants. Using the coefficients of several of weighted least square logit regressions for grouped data, we estimated the loan default rates for different definitions of entry and different levels of information about local market economic conditions. We found significantly higher default rates for banks that enter local markets than for incumbents. Default rates are higher for banks that enter without opening a branch, suggesting that having a branch on site can help to reduce the informational disadvantage. Our results confirm the insights provided by theoretical models, which emphasize the role of asymmetric information in determining the incentives and costs of entry into credit markets.

According to our estimates, entry into credit markets can generate substantial costs in terms of loan losses and therefore the number of entrants into local credit markets may be too low. Combined with the well-established negative correlation between the measure of market competition and loan interest rates (Berger and Hannan, 1989 and 1998), our results suggest that market free entry may not sweep away all of the monopolistic rents earned by incumbents.

References

Angelini, P. and Cetorelli, N. (2003) The effect of regulatory reform on competition in the banking industry, *Journal of Money, Credit and Banking* **35**, 663–684.

Berger, A. N., Kashyap, A. K., and Scalise, J. M. (1995) The transformation of the U.S. banking industry: What a long, strange trip it's been, *Brooking Papers on Economic Activity* **2**, 55–201.

Berger, A. N. and Hannan, T. H. (1989) Price-concentration relationship in banking, *Review of Economics and Statistics* **71**, 291–299.

Berger, A. N. and Hannan, T. H. (1998) The efficiency cost of market power in the banking industry: A test of 'quiet life' and related hypothesis, *Review of Economics and Statistics* **80**, 454–465.

Black, F. (1975) Bank funds management in an efficient market, *Journal of Financial Economics* **2**, 323–339.

Bonaccorsi di Patti, E. and Gobbi, G. (2001) The changing structure of local credit markets: Are small businesses special?, *Journal of Banking and Finance* **25**, 2209–2237.

Boot, A. W. A. and Thakor, A. V. (2000) Can relationship lending survive competition?, *Journal of Finance* **55**, 679–713.

Boyd, J. H. and De Nicolò, G. (2005) Bank risk taking and competition revisited, *Journal of Finance* **60**, 1329–1343.

Brander, J. A. and Lewis, T. R. (1986) Oligopoly and financial structure: The limited liability effect, *American Economic Review* **76**, 956–970.

Broecker, T. (1990) Creditworthiness tests and interbank competition, *Econometrica* **58**, 429–452.

Buch, C. M. (2002) Financial market integration in the U.S.: Lessons for Europe?, *Comparative Economic Studies* **44**, 46–71.

Cannari, L. and Signorini, L. F. (1997) Community links, co-operative rules, and the economic efficiency of Italy's local co-operative banks, in *Proceedings of the European Regional Science Association 37th Congress*, Rome, Italy.

Chan, Y., Greenbaum, S. I., and Thakor, A. V. (1986) Information reusability, competition and bank asset quality, *Journal of Banking and Finance* **10**, 243–253.

Danthine, J. P., Giavazzi, F., Vives, X., and Von Thadden, E. L. (1999) The future of European banking, *Monitoring European Integration*, Vol. 9, CEPR.

Degryse, H. and Ongena, S. (2004) The impact of bank competition on bank orientation and specialization, discussion paper 4543, CEPR.

Degryse H. and S. Ongena (2005) Distance, lending relationships, and competition, *Journal of Finance* **60**, 231–266.

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

Dewatripont, M. and Tirole, J. (1994) *The Prudential Regulation of Banks*, MIT Press, Cambridge MA.

DeYoung, R. and Hasan, I. (1998) The performance of *de novo* commercial banks: A profit efficiency approach, *Journal of Banking and Finance* **22**, 565–587.

Duan, N. (1983) Smearing estimate: A nonparametric retransformation method, *Journal of the American Statistical Association* **78**, 605–610.

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

Fama, E. F. (1985) What's different about banks?, *Journal of Monetary Economics* **15**, 29–40.

Focarelli, D. and Panetta, F. (2004) Are mergers beneficial to consumers? Evidence from the market for bank deposits, *American Economic Review* **93**, 1152-1172

Gehrig, T. (1998) Screening, cross-border banking, and the allocation of credit, *Research in Economics* **52**, 387–407.

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

Hannan, T. H. and Prager, R. A. (1998) The relaxation of entry barriers in the banking industry: An empirical investigation, *Journal of Financial Services Research* **14**, 171–188.

Jappelli, T., Pagano, M., and Bianco, M. (2005) Courts and banks: Effects of judicial enforcement on credit markets, *Journal of Money Credit and Banking* **37**, 223–244.

King, R. G. and Levine, R. (1993) Finance and growth: Schumpeter might be right, *Quarterly Journal of Economics* **108**, 717–737.

Manning, W. G. (1998) The logged dependent variable, heteroscedasticity, and the retransformation problem, *Journal of Health Economics* **17**, 283–295.

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

Mester, L. J., Nakamura, L. I., and Renault, M. (2005) Transaction accounts and bank monitoring, working paper 05-14, Federal Reserve Bank of Philadelphia.

Miller, M. J. (2003) Credit reporting systems around the globe: The state of the art in public credit registers and private credit reporting firms, in M. J. Miller (ed.), *Credit Reporting Systems and the International Economy*, Cambridge, MIT Press.

Petersen, M. and Rajan, R. G. (1995) The effect of credit market competition on lending relationships, *Quarterly Journal of Economics* **110**, 407–443.

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

Riordan, M. (1993) Competition and bank performance: A theoretical perspective, in C. Mayer and X. Vives (eds.), *Capital Markets and Financial Intermediation*, Cambridge University Press, Cambridge.

Sapienza, P. (2002) The effects of banking mergers on loan contracts, *Journal of Finance* **57**, 329–368

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

Sharpe, S. A. (1990) Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships, *Journal of Finance* **45**, 1069–1087.

Shleifer, A and Vishny, R. W. (1993) Corruption, *Quarterly Journal of Economics* **108**, 599–617.

Stein, J. C. (2002) Information production and capital allocation: Decentralized versus hierarchical firms, *Journal of Finance* **57**, 1891–1922.

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

Vives, X. (1991) Banking competition and European integration, in A. Giovannini and C. Mayer (eds.), *European Financial Integration*, Cambridge University Press, Cambridge.