Multilevel Empirics for Small Banks in Local Markets

Francesco Aiello[a], Graziella Bonanno[a,b]
f.aiello@unical.it - graziella.bonanno@unical.it

[a] Department of Economics, Statistics and Finance, University of Calabria
I-87036, Arcavacata di Rende (CS), Italy
[b] Royal Docks Business School, University of East London

Abstract Banking is increasingly a-spatial. However, the environment matters for small banks. Indeed, small banks are embedded in narrowed markets and hence benefit from proximity to their member-customers. Using a multilevel approach, this article measures how much the performance of Italian mutual-cooperative banks depends on both geographical and individual characteristics. After controlling for the time-effect, we estimate and distinguish the role of individual (small bank level) and local (provincial level) variables on banks' cost efficiency. Local markets effect explains 28.27% of bank heterogeneity in the basic empty multilevel model and 33% in the most extended model. Moreover, we find that bank efficiency increases with market concentration and demand density and decreases as bank branches increase in local markets.

JEL codes: G21, C13, D00, R19
Keywords: Multilevel model; mutual-cooperative banks; local markets; cost efficiency

1 Introduction

A large number of data in social sciences follows a hierarchical order. Returns to education depend on pupil's skills and efforts, but class also matters. Classes are further nested in schools. Again, firm performances are the result of individual behavior, although the environment in which they operate is crucial (Beugelsdijk 2007). In the field of industrial and regional economics, the hierarchical structure of micro-data has recently been handled by recurring to multilevel models (hereafter MLMs), which have been proved to perform well in estimating the spatial effect on firms productivity (Aiello et al. 2014, 2015; Fazio and Piacentino 2010; Van Oort et al. 2012), on firm attitude to cooperate in innovation (Srholec 2015) and in understanding the link between urbanization and firm innovativeness (Srholec 2010).

In banking, the recurrence of clustered data is more difficult, as this industry tends to be dominated by big-banks, whose organization is increasing complex and, very often, a-spatial. However, even in banking there is a phenomenon that can be modeled through MLMs. It regards the small-banks, which act as single-market entities. In other words, they are embedded in narrowed markets, representing, thereby, a good example of hierarchy: small-banks are at the lowest level of the hierarchy, while the higher level is the local market in which they act.

The embeddedness of small-banks in local markets makes easy to argue that environmental factors influence individual performance. This is not a novelty, as it is addressed by a massive literature exploring the bank performance-environment nexus.[1] Surprisingly, this literature refers to single-equation models that are too limited in handling the multilevel nature of data featuring small-banks behavior. On the contrary, the embeddedness may be properly treated

[1] See, above many others, Battaglia et al. (2010), Bos and Kool (2006), Dietsch and Lozano-Vivas (2000), Girardone et al. 2004). Hughes and Mester (2008) is a comprehensive survey on this topic.

by MLMs (Goldstein 2003; Luke 2004), which are very attractive also from an economic perspective, because they address how the micro, middle and macro spheres of economic systems evolve and interact. Indeed, if the first-level unity-of-analysis, the small bank, is embedded in a local market, then its performance cannot be addressed without taking into consideration the interactions from micro to macro level, and vice-versa, as multilevel does (Baldwin and Okubo 2006; Beugelsdijk 2007).[2] In this respect the hierarchical approach represents an important contribution for empirical studies aimed at understanding the individual performance and the links between micro and macro patterns (Aiello et al. 2014; Raspe and van Oort 2011; Srholec 2010, 2015).

These general considerations introduce to the methodological advantages of MLMs over single-equation framework. The latter disregards the nested structure of data, even though it is widespread known that ignoring this structure yields biased estimates of standard errors and subsequent increase in Type I error (Hox 2002).[3] Furthermore, MLMs combine different levels of data aggregation and relate them in ways that render the simultaneous existence of distinct level-one (small banks) and level-two (local markets) equations explicit. This allows the evaluation of whether, and to what extent, local factors matter in determining small-banks performance. In fact, on one hand the role of contextual factors is detected by testing hypothesis operating at different levels. On the other hand, MLMs decompose the heterogeneity in the output variable, providing highly informative outcome on “how much” contextual and individual factors explain of small-banks performance. Finally, compared to single-equation models, MLMs address the issue of error correlation across small-banks and the ecological and atomistic fallacies. (Heck and Thomas 2000; Hox 2002; Mass and Hox 2004).

As MLMs have never been used to study the role of context in banking, this paper tries to fill this gap. It refers to a MLM specification which is consistent with the hierarchical structure of data as well as with the fact that small-banks are observed over time. Thus, we refer to a MLM for longitudinal panel, given that multiple measurement at different time points (level 1) are nested within small-banks (level 2), which are further nested in local markets (level 3). In order to represent time changes, a growth linear MLM with a random intercept as well as random slopes is considered. Furthermore, the analysis is expanded by including a set of predictors at each level of our hierarchy that we retrieve from the banking literature.

Besides the method, another element of originality of this paper comes from the analyzed case-study, regarding the Italian mutual-cooperatives banks (hereafter MCBs). This is an interesting case because in Italy, as well as in many other countries, banking has been strongly deregulated over the last two decades.[4]

[2] The links between agents and external factors are modeled from different perspectives. For instance, the endogenous growth theory proves the existence of increasing returns due to spillovers between firms and other higher level organizations (Romer 1986). However, it refers to uni-equational macro models and focuses on aggregate patterns, although they have micro-foundations. Again, the existence of micro-macro interactions is also recognized by the evolutionist school. However, here the links are one-way, in the sense that they flow from the individual to the aggregate level (Dosi and Nelson 2010). This implies that the “overall” patterns are just those from aggregations, while any other important environmental factor is left out from the analysis.

[3] In MLMs the inference is made by distinguishing between sample size at the different levels of data aggregation. One consequence of failing to recognize hierarchical structures is that standard errors of regression coefficients will be underestimated, leading to an overstatement of statistical significance.

[4] The major reforms go back to the 1990s' when the 1990 Amato-Carli Act, the EU Directive II and the 1993 Consolidated Act. During the successive decade, the 2002 budget law, the 262/2005 law and the

An important reform was the relaxing of geographic constraints, allowing banks to open branches wherever, thereby inducing a territorial diversity in banking and more competition even in the periphery. Notwithstanding these profound changes, the permanence of a plethora of small-banks appears a paradox, as MCBs survival is threatened by two main forces. Firstly, the increased action of complex financial conglomerates forces the disappearance of small entities: in a world of big-banks, small credit institutions are expected to disappear. Secondly, MCBs historically operated in narrow isolated markets, which, now, are no longer protected with regulatory barriers. If local markets become contestable, then it is expected that MCBs will lose their quasi-monopoly power which, in the past, assured profitability (Coccorese 2009; Fiordelisi and Mare 2013) Differently from expectations, MCBs reacted to reforms by re-organizing their network through within-group M&A and thus preserving their presence in local markets (Gutiérrez 2008).[5] Finally, the interest in evaluating the role of local market conditions on MCBs performance is amplified by the fact that the banking sector remains highly heterogeneous in Italy, with marked territorial differences throughout the country.

Covering the 2006-2011 period, this paper uses micro-data from the "Italian Banking Association" and provincial data from Bank of Italy and ISTAT (the Italian Institute of Statistics). MCBs performance is measured by the cost efficiency scores, which have been estimated through the stochastic frontier approach.

The main results are as follows. After having found that heterogeneity in MCBs performance exists at the beginning of the period and in the trajectory of change over time, we show that the differences in MCBs internal characteristics matter. However, location across Italian provinces plays a relevant role in explaining MCBs behavior. To be more precise, in the most extended MLM specification 31% of the variance in MCBs' cost efficiency is due to banks characteristics and 33% is ascribable to location, while the remaining 36% is due to the time-effect over the turbulent years-crisis under scrutiny. We also find significant links between MCBs efficiency and a number of variables (i.e., market concentration, branching, demand density) capturing the local market conditions.

The paper is organized as follows. Section 2 presents models and data. Section 3 discusses the results and section 4 concludes.

353/2006 Legislative Decree speed up the processes of consolidation and market competition. Details are in Giannola (2009), Messori et al. (2003) and Silipo (2009).

[5] At the end of 2013 there were 385 (411 in 2011) MCBs, while in the early 90s' they were 700. However, a consolidation process in the network occurred involving mostly MCBs, with the result that the number of branches even doubled in ten years, moving from 2226 in 1993 to 4454 in 2013. In relative terms, in 2013 MCBs branches made up 14% of total national branches, which is a value 4 percentage points higher than that of 1993.

2 The empirical setting: model, data and variables

2.1 The model

Understanding whether and how market conditions affect small banks performance is a typical example of hierarchy, in the sense that the units (small banks) refer to different levels of aggregation (local markets) (Goldstein 2003). If a nested structure of data exists, single-level methodologies will suffer from some estimation problems. First, as a result of locally specific factors, small banks operating in a market are likely to be more similar than small banks located in differing areas, implying that residuals are not independent. This issue is addressed by the multilevel approach which, controlling for territorial-effects, ensures more efficient estimates than uni-equational model. In addition, single-level regressions yield an inflated significance of level-two coefficients because the diagnostics refers to the number of level-one observations instead of the number of higher-level units. It is likely that the significant relationships found in OLS regressions will turn out not to be significant in multilevel regressions. In other words, the multilevel model controls for spatial dependence and corrects the measurement of standard errors, so reducing the risk of type I errors.

Apart from the statistical improvements, another advantage of the multilevel model is that variables at different levels are not simply add-ons to the same single-level equation, but are linked together in ways that make the simultaneous existence of distinct level-one and level-two equations explicit. In such a way, level-two factors are used not just as independent variables to explain variability in a level-one dependent variable, but also to explain variability in random intercept and random slopes (Bickel 2007). However, in order to explain the variability in random coefficients, a "sufficient" number of clusters in the sample is required. Otherwise the between-group variance is poorly estimated. In this respect, a clear result does not exist (Richter 2006), although there are some rules of thumb, which, however, are very different from each other.[6] In our empirical setting, this issue is absent as the number of groups are large enough to ensure the reliability of results.

In this study, MCBs are observed over the 2006-2011 period. The six time points are the first level of the analysis, while MCBs and local markets constitute, respectively, the second and the third levels of our hierarchy. The three-level units refer to local banking markets that we identify with the Italian provinces (NUTS 3 codes) whose initial number is 103. As in some provinces the number of MCBs is limited (even zero), the analysis is performed by considering 66 provinces with more than two small banks per year. On the other hand, the number of small-banks is high in other markets (see appendix Table A1). This heterogeneity in cluster-size is the motivation behind the choice to use a three-level model for longitudinal data with randomness in intercepts and slopes. Since data follow a longitudinal structure, the MLM specification treats time as source of randomness both in the intercepts and slopes at any level. The dependent variable y refers to MCBs (second-level) in a specific time period (first-level) and depends on a set, say MCB, of variables measured at individual bank level and on a set, say P, of variables defined at provincial

[6] Some authors suggest 20 groups (Heck and Thomas 2000; Rabe-Hasketh and Skondal 2008), others 30 (Hox 2002) or 50 (Mass and Hox 2004). In addition, it is worth noting that the clusters must be sized with at least two observations in random-effects models. In the random effect specification, smaller groups have a smaller impact on the estimation results than larger groups (Snijders and Berkof 2008). This approach recognises that there is little information for small groups by "shrinking" their residual estimates towards zero and, therefore, pulling their mean towards the overall mean (Bickel 2007).

level (third level of the hierarchy). The dependent variable may be predicted as follows (level-one model):

$$y_{tij} = \beta_{0ij} + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + \delta_{.ij} Time + e_{tij} \qquad [1]$$

where y_{tij} is the vector of the estimated MCBs cost efficiency; β_{0ij} is the intercept, β_{1ij} are the slope coefficients and e_{tij} is the random error term with zero mean and variance σ_e^2; $\delta_{.ij}$ is the slope associated to the time variable; *t* is for time (*t=2006…2011*), *j* is for province (*j=1…p*) and *i* states for MCB in province j (*i=1…N*$_j$). The error term e_{tij} captures not only residual variance, as OLS regression does, but also potential group-to-group variability in the random intercepts and slopes. In eq.[1] e_{tij} are not independently distributed, because of nesting: MCBs operating in the same province tend to have correlated residuals, so violating the assumption of independence.

The parameters β_{0ij} and $\delta_{.ij}$ of eq. [1] vary across banks and provinces. At MCBs level they are modelled as:

$$\left.\begin{aligned} \beta_{0ij} &= \gamma_{00j} + u_{0ij} \\ \delta_{.ij} &= \delta_{.0j} + u_{.ij} \end{aligned}\right\} \text{level} - \text{two model} \qquad [2]$$

Substituting eq.[2] in eq.[1] yields:

$$\begin{aligned} y_{tij} &= \beta_{0ij} + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + \delta_{.ij} Time + e_{tij} \\ &= (\gamma_{00j} + u_{0ij}) + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + (\delta_{.0j} + u_{.ij}) Time + e_{tij} \\ &= \gamma_{00j} + u_{0ij} + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + \delta_{.0j} Time + u_{.ij} Time + e_{tij} \end{aligned} \qquad [3]$$

Similarly, γ_{00j} and $\delta_{.0j}$ of eq. [3] may be expressed as:

$$\left.\begin{aligned} \gamma_{00j} &= \gamma_{000} + u_{00j} \\ \delta_{.0j} &= \delta_{.00} + u_{.0j} \end{aligned}\right\} \text{level - three model} \qquad [4]$$

Substituting eq.[4] in eq.[3] and after algebra manipulations one gets the full mixed model:

$$y_{tij} = \gamma_{000} + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + \delta_{.00} Time + u_{00j} + u_{0ij} + u_{.0j} Time + u_{.ij} Time + e_{tij} \qquad [5]$$

where γ_{000} is the overall mean, u_{0ij} is random departure from the overall mean due to the i-*th* MCB, u_{00j} is random departure from the overall mean due to the j-*th* province, $u_{.ij}$ captures the departure from the common linear trend due to the i-*th* MCB and $u_{.0j}$ is the departure from the common linear trend due to the j-*th* province. Finally, $e_{tij} \sim N(0, \sigma_e)$ represents the deviation due to time effect.

The econometric model [5] is composed by a deterministic part - $\gamma_{000} + \beta_1 MCB_{tij} + \beta_2 P_{t.j} + \delta_{.00} Time$ - which contains all the fixed coefficients - and by a stochastic component - which is represented by *u*-terms and e_{tij}. Besides e_{tij}, the stochastic part is the sum of two components: the $u_{00j} + u_{.0j} Time$ is the random part associated to the level-three of the model, while $u_{0ij} + u_{.ij} Time$ is related to the level-two model

An important aspect of MLMs is that they allow to decompose the MCBs heterogeneity in efficiency to the contribution of unobserved factors at any level of data aggregation. To this end, let's consider the "empty" specification of eq. [5], i.e. the model without time and explanatory variables:

$$y_{tij} = \gamma_{000} + u_{00j} + u_{0ij} + e_{tij} \qquad [6]$$

Eq. [6] allows the decomposition of the unobserved variance of y into three independent components, i.e. the variance of e_{tij} (σ_e^2), the so-called within-group variance, the variance of u_{00j} (σ_{uj}^2), also known as between-group variance for provinces and the variance of u_{0ij} (σ_{ui}^2), which is the between-group variance for MCB-level.

A useful index to evaluate the relative magnitude of the variance components is the intra-class correlation (ICC). It measures the proportion of the response variance that lies at each level of the hierarchy. ICCs is calculated level-by-level and differ model-by-model. For instance, as far as the provinces are concerned, the ICC is given by the ratio of the variance at that level, σ_{ui}^2, to the total variance, that is:

$$ICC_j = \frac{\sigma_{\mu j}^2}{\sigma_{\mu j}^2 + \sigma_{\mu i}^2 + \sigma_e^2} \qquad [7]$$

Similarly, the ICCs for MCB and time level are, respectively:

$$ICC_i = \frac{\sigma_{\mu i}^2}{\sigma_{\mu j}^2 + \sigma_{\mu i}^2 + \sigma_e^2} \qquad [8]$$

$$ICC_t = \frac{\sigma_e^2}{\sigma_{\mu j}^2 + \sigma_{\mu i}^2 + \sigma_e^2} \qquad [9]$$

Of course when considering the full mixed model (eq. 5), the ICCs take into account the entire structure of variance, as randomness also entails time-slopes. Put differently, eq. [7]-[9] refer to a general formulation of MLM model, in the sense that comprise all the alternatives arising from the combination of the sources of randomness. For instance, for eq. [5] we have that $\sigma_{\mu j}^2 = \sigma_{\mu j}^2\_\text{intercept} + \sigma_{\mu j}^2\_\text{slope}$ and $\sigma_{\mu i}^2 = \sigma_{\mu i}^2\_\text{intercept} + \sigma_{\mu i}^2\_\text{slope}$. At the opposite side, when estimating a random intercept model (for both MCBs and provinces levels as in eq. [6]), the variance is given by $\sigma_{\mu j}^2 = \sigma_{\mu j}^2\_\text{intercept}$ and $\sigma_{\mu i}^2 = \sigma_{\mu i}^2\_\text{intercept}$. In between these two extremes there are other MLM specifications, depending on whether modelling the slopes randomness due to time.

2.2 Data at bank level

Microdata are from the Italian Banking Association (ABI) which collects balance-sheets of about 97% of Italian banks. This dataset contains information of more than 400 MCBs per year. After cleaning data procedure, we use an unbalanced panel of 2334 observations.[7] On average they are more than 63% of the sample [the remaining are corporations (32%) and Popolari banks (6%)]. It is noteworthy to point out that, over the 2006-2011 period, MCBs size is, on average, 295M Euro (Table 1), that is to say about thirty times smaller than the size of other banks (6,903M Euro).

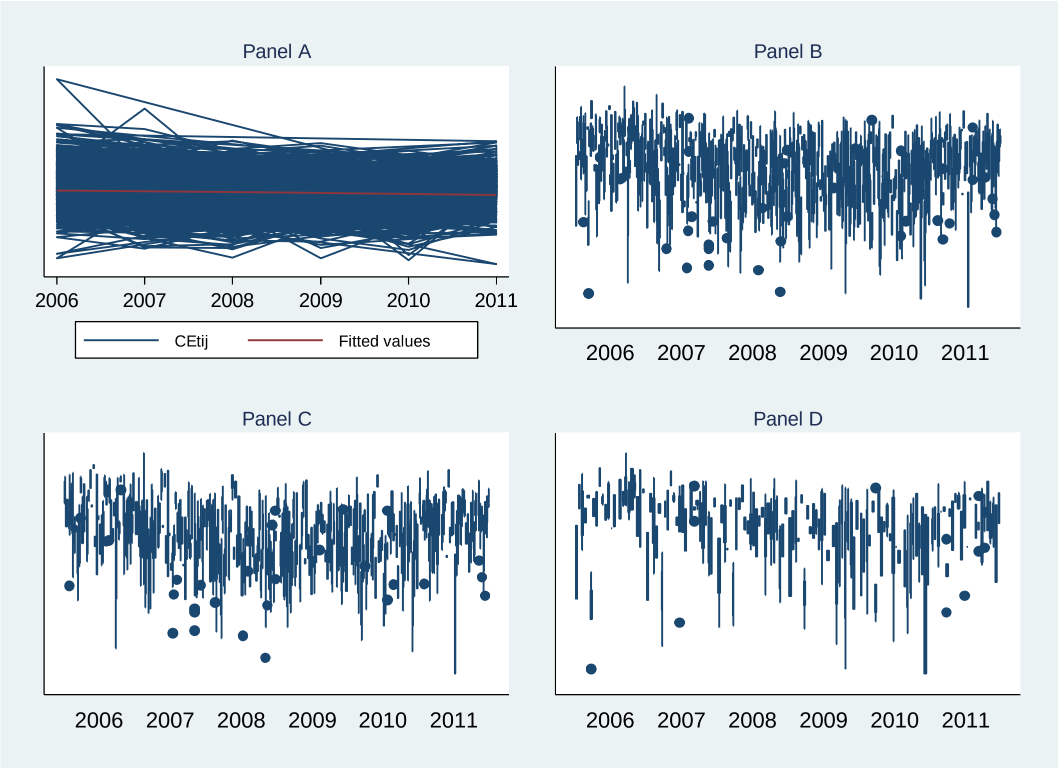
Table 1 highlights that cost efficiency of MCBs is 0.80, thereby meaning that MCBs should reduce the inputs of 20% offering the same banking services (or similarly they should increase outputs of 20% with the same inputs). Data indicate that cost efficiency is quite dispersed (the minimum value is 0.24 and the maximum is 0.98). Its dynamics over the six year-points 2006-2011 is displayed in Figure 1 (Panel A). It is clear that there is a considerable inter-MCBs heterogeneity at the beginning year 2006 and a high variation over time. The red line refer to the intercept and growth for the whole sample of MCBs. Figure 1 also plots the MCBs cost efficiency by province and year. Panels B-D of Figure 1 indicate that the within and between group heterogeneity is high across all province (Panel B), whichever the MCBs location in Northern (Panel C) or in Southern provinces (Panel D). These visual representations of cost-efficiency further legitimate the use of MLM for longitudinal data, whose estimations provides a statistical test of the variability in intercepts and growth terms - as depicted in Figure 1.a - and of the role of any hierarchical level of data in explaining the variability of individual outputs.

Turning back to Table 1, when referring to the cost income the average is 0.73, with a minimum at 0.4 and the maximum at 3.53. The analysis of other individual profiles reveals that MCBs activities are weakly diversified in terms of income or loans diversification.[8] Income diversification is 0.21, while MCBs loan diversification is 0.32. MCBs ability to transform deposits into loans is, on average, 1.51. Interestingly, the ratio Equity/Total Assets is significantly low (0.018), thereby meaning that MCBs show high financial dependence.

[7] The size of our panel is data driven and corresponds to the number of the estimated values of cost efficiency, which is the dependent variable used in our multilevel models. We drop from the MCBs with missing (or zero-) values of total assets and employees. Cost efficiency estimations are from a stochastic translog frontier in the specification proposed by Battese and Coelli (1995). The cost equation is a 3-inputs-3-outputs model, while the inefficiency equation only controls for bank-type (MCBs, Popolari and Ltd) and location effects. This has two advantages. Firstly, estimations refer to a national cost frontier, ensuring comparability of the estimated efficiency, as MCBs performance is relative to the rest of the industry. This allows to control for between-groups effect. Secondly, the estimated efficiency scores are net of any institutional and geographical effect. Results from stochastic frontier estimations are available upon request.

[8] Income diversification is calculated as [Income Commissions /(Income Commissions + Net Interests Income)], while loans diversification is (1-Loans/Total Assets).

Figure 1
MCBs cost efficiency by time and province (2006-2011)



2.2 Data at provincial level

The estimation of a MLN requires a set of variables capturing the local market conditions. As already mentioned, this paper refers to the province (NUTS3 code) as reference market of MCBs. An analysis based on larger territories, for example regions (NUTS2) as in Battaglia et al (2010), could suffer from aggregation bias. Phrased differently, it is plausible that the greater the proximity of MCBs to markets the more precise will be the investigation of the individual efficiency-environment nexus. Thus, using provinces as territorial unit-of-analysis assures that MCBs performance may be intended as the result of banking relationships between MCBs and the "residents".

Said this, this section documents some characteristics of banking markets across 103 out of 110 Italian provinces.[9] Data are from Bank of Italy. An important effect of the restructuring reform is the spatial diffusion of financial services. Several proxies can be used as an indicator of this. For instance, the bank branches by square kilometer measures the density by province. It is on average 0.0014, with considerable variation across provinces (it varies from 0.0002 for Crotone to 0.0129 for Milan in 2006-2011). An additional indicator is the ratio "Bank Branches/Municipalities" per province, which is, on average, more than 5 in 2006-2011 (it ranges from more than 20 branches per municipality in the provinces of Trieste and Prato to less than one branch per municipality in the provinces of Isernia, Oristano and Vibo Valentia). Along this line of reasoning, further evidence comes from the concentration of provincial markets. The Hirschman-Herfindahl

[9] Including information of new 7 provinces undermines data comparability over time.

index calculated using, by year, the number of branches per bank (HH1) in every province is 0.125 over the years 2006-2011, falling in the range 0.036-0.537. Higher average market concentration has been revealed when considering total bank assets (HH2).[10] In this case, the HH2 index averaged by year and province is 0.208 (about two times higher than the average of HH1) varying from 0.030 and 0.506. Finally, there has been a relevant increase of big-bank participation in the periphery. The top-3 national banks - as revealed by the total assets averaged over 2006-2011 - owned 21% of bank branches operating in every Italian province (Table 1). [11]

Another issue concerns the traditional function of banks, namely the transformation of deposits into loans. The Bank of Italy provides the required data taking into account the residence of customers and depositors. High values of this ratio mean that the provincial banking sector is issuing out more of its deposits in loans at provincial level, which, in turn, means it releases more income. Over 2006-2011 the provincial ratio Loans/Deposits is on average 1.548. The highest value (3.046) is in Milan, whilst the lowest (0.729) refer to the province of Trieste. A related issue to offering funds is that loans are not always repaid. In Italy, bad performing loans are 6.38% of total loans in 2006-2011, with a different incidence across provinces. In some provinces (Milan, Sondrio and Siena), bad-loans are low (less than 2%), while they are very high (more than 10%) in Avellino, Benevento, Caserta, Crotone, Caltanissetta, Enna, Frosinone, Isernia, Latina, Nuoro, Potenza, Reggio di Calabria, Taranto, Vibo Valentia and with a peak of 18.45% in Matera. Finally, there is also great heterogeneity when looking at the credit provided by banks: the loans-to-GDP ratio which ranges from the highest value of Milan (3.454) to the lowest values (0.392) of Vibo Valentia. It is interesting to point out that data of Table 1 reproduce the North-South dualism of the Italian economy. Indeed, in the South of the country there is less access to banking services, more concentration of bank branches, low financial development and loans/deposits ratio, while the incidence of non-performing loans is high (Table 1).

From the above discussion one learns that the local banking market conditions are still extremely heterogeneous across Italian provinces. This marked heterogeneity further motivates the understanding of the nexus between local determinants and MCBs performance.

---

[10] Data needed to calculate HH2 is the value of total assets by the i-*th* bank in every province j (TAij). Because this information is not freely available in Italy, as well as in many other countries, we proceed through this calculation: $TA_{ij}=TA_i*b_{ij}$, where TAi is the balance-sheet amount of Total Asset (TA) of the i-th bank and bij is the proportion of branches of bank i in province j (bij=BBij/BBj). This procedure is proposed by Carbò Valverde et al. (2003).

[11] The role of big-banks in local markets is more apparent when looking at their total assets shares. The top-3 banks absorbed (on average) 73% of total assets at provincial level in 2006-2011 (table 1). The territorial distribution of this market share shows a minimum of 51% in Benevento and a maximum in Siena (more than 90%). It is worth pointing out that in 22 out of 103 Italian provinces, the top-3 national banks absorb more than 80% of local total assets Alessandria, Aosta, Como, Imperia, Mantova, Milan, Novara, Pavia, Turin, Belluno, Arezzo, Grosseto, Massa, Siena, Lecce, Agrigento, Caltanissetta, Enna, Messina, Ragusa, Syracuse, Trapani (detailed statistics for each province are available upon request).

Table 1
Overview of data at bank and provincial level over the 2006-2011 period

	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Data at bank level					
Size	2334	295.30	376.22	4.4400	5465.35
Cost Efficiency	2334	0.8034	0.0706	0.2426	0.9779
Cost Income	2334	0.7267	0.1662	0.4049	3.5340
Income Diversification	2334	0.2102	0.0669	0.0013	0.7102
Loans Diversification	2334	0.3207	0.1277	0.0959	0.9758
Loans/Deposits	2334	1.5080	0.5759	0.0650	4.0582
Equity/Total Assets	2334	0.0175	0.0533	0.0001	1.1972
Data at provincial level (All Italian provinces)					
Branches by square Km	103	0.0014	0.0016	0.0002	0.0129
Branches per municipality	103	5.2915	4.5120	0.6630	23.78
Market concentration on bank branches (HH1)	103	0.1245	0.0701	0.0360	0.5370
Market concentration on total assets (HH2)	103	0.2084	0.0902	0.0300	0.5060
Share of the top 3 banks (by Total Assets)	103	0.7265	0.0904	0.5063	0.9058
Share of the top 3 banks (by bank branches)	103	0.2083	0.0901	0.0297	0.5055
Loans/deposits	103	1.5477	0.4603	0.7290	3.0460
Non performing loans	103	6.3806	3.4688	1.2700	18.45
Financial development (Loans/GDP)	103	0.9681	0.4423	0.3920	3.4540
Data at provincial level (Southern Italian provinces)					
Branches by square Km	34	0.0007	0.0011	0.0002	0.0070
Branches per municipality	34	3.5478	2.8884	0.6630	12.1670
Market concentration on bank branches (HH1)	34	0.1489	0.1037	0.0750	0.5370
Market concentration on total assets (HH2)	34	0.1959	0.0851	0.0760	0.3340
Share of the top 3 banks (by Total Assets)	34	0.7045	0.1001	0.5063	0.8610
Share of the top 3 banks (by bank branches)	34	0.1958	0.0851	0.0762	0.3342
Loans/deposits	34	1.1743	0.2169	0.8800	1.7390
Non performing loans	34	9.9692	2.9201	5.1860	18.4490
Financial development (Loans/GDP)	34	0.6363	0.1718	0.3920	1.1330

Source: our elaborations on data from ABI and Bank of Italy

3 Heterogeneity in MCBs performance: the empty MLM and the time-effect

This section refers to the estimations obtained when considering different MLN specification, ranging from the MLNs incorporating only time (that is to say with $\beta_1 = \beta_2 = 0$ in eq. [5]) to the empty model (eq. [6]). Since we use longitudinal data, our hierarchy is composed by three levels: time (first level), MCBs (second level) and provinces (third level). Table 2 displays the MLMs results for cost efficiency regressions.[12] Column 1 of Table 2 refers to the random-intercept empty model in which the second level is formed by 414 MCBs and the third level by 66 provinces. Observations are 2334. In column 2, time enters into the deterministic part of the model to depict growth. Columns 3-5 refer to the estimations adding randomness in the second and/or third level slopes. Important diagnostic to choose the best performing regression comes from the AIC test.

The first outcome to be discussed is the likelihood-ratio test, which compares the MLN with the standard OLS regression. If the null hypothesis is true, OLS can be used instead of a variance-components model. The test result of column 1 supports the use of multilevel methodology and indicates that the intercept should be considered as a group-by-group variant coefficient. The evidence in favor of the multilevel approach holds for each model considered in Table 2, thereby supporting the conclusion that Italian MCBs behavior follows a hierarchical structure.

Before discussing the role of individual and local unaccounted heterogeneity, it is remarkable to highlight that the coefficient of Time is always negative, indicating that during the years of the current crisis the MCBs register significant losses in cost efficiency. This estimate confirms the pattern of the red line in Figure 1. However, the nexus between crisis and small-banks efficiency deserves to be investigated better, as made by Barra et al. (2014).

As can be seen from the estimations of the empty model (column 1 of Table 2), the province-specific unobservable factors capture 28.27% of the MCBs heterogeneity in efficiency, while the remaining is explained by MCBs (28.11%) and time (48.11%) effects. Moving from one model to another, the portion of variance explained by each level varies a lot.[13] For instance, the ICC index of the provinces is high (40.05%) when time enters as source of randomness of provincial intercepts and slopes (Models 4 and 5), while the role of unaccounted MCBs factors remains broadly the same, falling in the range 21.63% of Models 4 and 5 and 23.8% of Model 2. In order to provide robustness to the evidence that the unobservable-province factors help a lot to detect the heterogeneity of MCBs efficiency, we re-run MLM regression year-by-year. In such case, the time-level disappears and the hierarchy is at 2-level - MCBs and provinces - implying that estimations refer to a random-intercept model, where disturbances from the overall-MCBs mean are just due to location. As the ICC for province is extremely high (ranging from 45.5% in 2006 to 30.2% in 2011), results of Table 3 firmly confirm that the environment in which small-banks operate is a key dimension to be taken into account when explaining their individual performance.

---

[12] In running MLM regressions, the dependent variable - that is the MCBs cost efficiency - has been transformed using the following formula: CE^{Trans}=ln(CE/(1-CE). This is because CE is never zero or unity, thereby making inappropriate the use of a Tobit model, which, on the contrary, performs well only if the upper and lower bounds come from non-observability (Maddala 1991; McDonald 2009).

[13] The discussion model-by-model points out (i) the different impact of time in the longitudinal setting we propose and (ii) the varying role of each hierarchical level in explaining individual outcome. However, the AIC assumes low values in Model 4 and 5 (about 1850) against the high values (from 1953 to 1960) of Models 1-3, suggesting that the best fitting refer to MLMs with time randomness in both MCBs/provincial intercepts and slopes.

Table 2

Explaining heterogeneity in cost efficiency of Italian MCBs over the 2006-2011 period. Results from the empty model and MLMs with intercept and slope time randomness.

	Dep. Variable: Cost Efficiency				
	No Time Effect	Time Effect			
		Intercepts	Intercepts and II level slopes	Intercepts and III level slopes	Intercepts, II and III level slopes
	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	1.469	1.525	1.531	1.595	1.595
	(42.82)	(41.35)	(40.92)	(36.26)	(36.26)
Time		-0.016	-0.016	-0.036	-0.036
		(-4.13)	(-4.03)	(-5.42)	(-5.42)
Random-Effects					
Variance					
Provinces (Intercept)	0.0600	0.0599	0.0631	0.0943	0.0943
Provinces (Slope)				0.0016	0.0016
MCBs (Intercept)	0.0502	0.0503	0.0473	0.0518	0.0518
MCBs (Slope)			0.0005		1.62E-25
Time Random Effect	0.1022	0.1013	0.0991	0.0918	0.0918
Total	0.2124	0.2115	0.2099	0.2396	0.2396
ICC					
Provinces	28.27%	28.31%	30.08%	40.05%	40.05%
MCBs	23.62%	23.80%	22.73%	21.63%	21.63%
Time	48.11%	47.89%	47.19%	38.32%	38.32%
LR test (p-value)	0.000	0.000	0.000	0.000	0.000
AIC	1960.09	1954.361	1953.77	1849.90	1851.90
Observations	2334	2334	2334	2334	2334
N. of Groups:					
MCB level	414	414	414	414	414
Province Level	66	66	66	66	66

Legend: z-value are in brackets.

LR test is for the choice between ML and linear regression (H0).

AIC=-2*Log-lik+2*k, where k is the number of estimated parameters.

Source: see Table 1

Table 3
Explaining heterogeneity in cost efficiency of Italian MCBs performance.
Evidence from a random-intercept multilevel model by year

	Dep. variable: Cost Efficiency					
Explanatory Variables	2006	2007	2008	2009	2010	2011
Constant	1.747	1.498	1.348	1.358	1.303	1.569
	(37.87)	(33.545)	(32.52)	(38.23)	(36.58)	(48.32)
Random-Effects						
Variance						
Provinces	0.104	0.093	0.083	0.052	0.051	0.047
MCBs	0.124	0.143	0.124	0.120	0.121	0.084
ICC provinces	45.5%	39.4%	40.0%	30.2%	29.5%	35.8%
LR test	88.82	82.32	83.15	62.51	66.47	86.21
p-value	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
Number of Groups	65	65	66	65	65	65
Number of observations	377	392	410	394	378	383

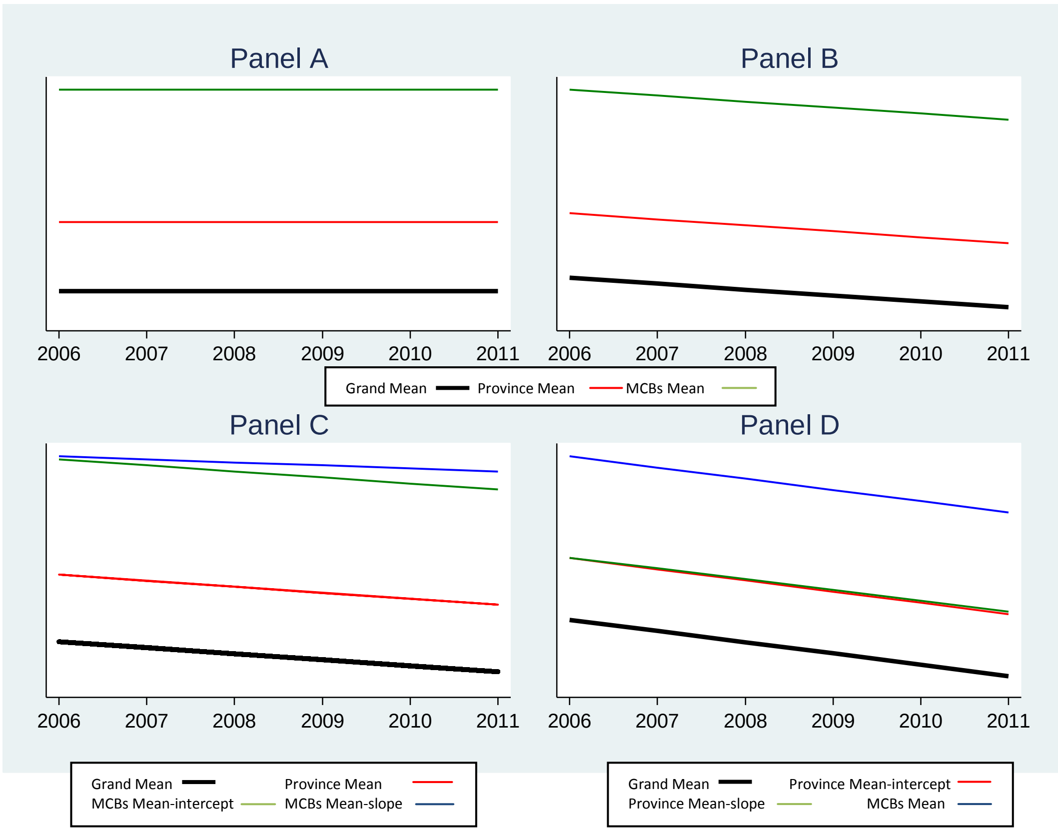
Legend: z-value are in brackets.

Source: see Table 1

Figure 2 depicts the results of Table 2. It also helps the understanding on how MLMs for longitudinal data actually work for the sample of Italian MCBs. The black line states for the overall mean γ_{000} of our MLM, that is the constant term in Model 1-4 of Table 2. In Panels A and B, the green line captures the mean at MCBs level, while the red line refers to the provincial mean. Panel A of Figure 2 reproduces Model 1 of Table 2, where nothing in the fixed or random parts of the model is a function of time. When considering time as part of the fixed component of the model (Model 2 of Table 2), the overall grand mean more accurately reflects the pattern of cost-efficiency over time. The negative sign of the time slope (-0.016 in Model 2) now translates into decreasing lines of the average effect at any level (Panel B). Finally, if Time enters into the random parts of the model, estimations will yield a separate regression line within each province or MCBs (Panels C and D).

From the above discussion we can argue that MCBs heterogeneity in performance is highly sensitive to individual factors and location. Given this and in order to explain better the role of province as a source of variability, we complement the analysis by augmenting the MLMs with some MCBs and provincial observables.

Figure 2
MCBs cost efficiency. MLN results in a graph



4. Augmenting the multilevel model with MCBs and provincial-specific variables

This section presents the results obtained when the MLM is augmented through a set of individual and provincial variables. Starting from a specification in which time is treated as source of randomness at any level, the aim of this section is twofold. Firstly, the evidence so far presented indicates that the proportion of efficiency variability explained by unobservable specific-effect is high. Therefore, after considering a selected set of determinants of efficiency, we expect to grasp part of this black-box of unaccounted individual heterogeneity. Secondly, our main interest remains in understanding the role of location, net of the role exerted by observables.

Regressors inserted in eq. [5] are size, diversification and the equity/total assets ratio at MCBs level, whilst at provincial level they concern the banking market concentration, the demand of banking service, the branching process and the riskiness of local banking markets. They have been already presented (*cfr*. Table 1). Moreover, in order to control for macro-economic effects, we also include the provincial GDP per-capita (sourced by ISTAT).

While results of Table 4 refer to MLM regressions for the entire sample of MCBs and provinces, Table 5 displays the estimates obtained when performing a sensitivity analysis. Table 4 follows the presentation of Table 1, whereas Table 5 uses the full-specification of the mixed-model (that is to say the one with the lowest AIC of Table 4). The sensitivity analysis of Table 5 is performed by splitting the sample according to the (i) banks location (northern and southern provinces in columns 1 and 2, respectively); (ii) MCBs cost efficiency distribution (1st quartile in column 3, 2nd and 3rd quartiles in column 4 and 4th quartile in column 5); (iii) MCBs size distribution in columns 6-8 (dividing the sample by using three areas of size distribution, as made for efficiency).[14]

Before discussing the role of observables, it is meaningful to highlight that the multilevel approach allows the possibility to calculate the coefficient of determination and obtain, at any level of the hierarchy, a proportional reduction in the estimated total residual variance when moving from the "empty model" to an extended specification of the model (Rabe-Hesketh and Skrondal 2008).[15] The overall fit of the Model 5 in Table 4 is 31.69% and is the result of different contribution at each level. On one hand, individual-MCBs variables absorb 9.8% of the variance estimated at the 2-level of the hierarchy. On the other hand, the R^2 at provincial-level is 20%. Interestingly, the set of observables at provincial level used in Table 4 contributes to explain always more than 20% of efficiency variability that we observe at that level, with a peak of 40% in Model 2. Table 5 points out that the goodness of fit differs a lot according to the sub-sample of MCBs we refer to. Finally, it is noteworthy to say that the observables do not impact on the relative values of ICCs. Data of Table 4 show that the proportion of MCBs heterogeneity in efficiency explained by location effect remains high, falling in the range between 24.22% (Model 2) and 33.08% (Model 5).

In what follows we summarize econometric results by firstly focusing on the individual factors and then discussing the evidence at provincial level.

At bank level, the first relevant issue regards the efficiency-size nexus. Regressions include the SIZE, expressed as the logarithm of total assets of each MCB. There is no prior expectation on the sign of the size-effect, as it may be positive (Andries 2011; Drake 2001) or negative (Pilloff 1996). We find that MCBs cost efficiency tends to decrease with size. However, as size enters into regressions in logarithmic terms, its marginal effect is nonlinear and tends to zero as size increases This is an interesting outcome, as it implies that the sensitivity of MCBs efficiency to size is extremely low above a certain threshold. For instance, if MCBs average size increased by 10%

---

[14] We replicate Tables 4 and 5 by addressing the issue of missing values in MLN for longitudinal data (Kwok et al. 2008; Little and Rubin 2002). To this end we employ the Stata command "mi impute" developed by Royston (2007; 2009). The missing are 221 over the 2006-2011 period, that is less than 10% of the entire sample. However, including missing values into regressions does not impact on the sign and significance of results displayed in Tables 4 and 5, that is those obtained without using the "mi impute" procedure (findings with missing are available upon request).

[15] The total pseudo-R^2 for the three-level model is given by:

$$R^2 = \frac{(\sigma^2_{\mu jN} + \sigma^2_{\mu iN} + \sigma^2_{eN}) - (\sigma^2_{\mu jM} + \sigma^2_{\mu iM} + \sigma^2_{eM})}{\sigma^2_{\mu jN} + \sigma^2_{\mu iN} + \sigma^2_{eN}}$$

where N stands for the empty model and M for the model of interest. The proportional reduction in the variance explained by observables can be calculated level-by-level. For instance the proportion of the level-3 variance explained by the covariates is: $R_3^2 = (\sigma^2_{jN} - \sigma^2_{jM}) / \sigma^2_{jN}$; the proportion of the level-2 variance explained by the covariates is: $R_2^2 = (\sigma^2_{iN} - \sigma^2_{iM}) / \sigma^2_{iN}$; and the proportion of the level-1 variance explained by the covariates is: $R_1^2 = (\sigma^2_{eN} - \sigma^2_{eM}) / \sigma^2_{eN}$

(passing from the actual 295MLM of euro to 324Mld of euro) the efficiency would diminish by 0.70%, passing, on average from 0.812 to 0.807 (calculations are based on Model 5 in Table 4, because of its lowest AIC). Importantly, the negative link between size and cost efficiency holds true whatever the sample we consider in Table 5.

There are some reasons to consider as relevant the diversification of activities for bank efficiency. It is argued that income from traditional bank activities suffers lower volatility than other financial uses and then the higher the share of risky activities the lower the exposure to systematic risk (Vallascas and Keases 2012). However, it is not certain that the higher betas coming from diversification compensate the costs for diversifying the sources of income (Baele et al. 2007; Wagner 2010). Therefore, there is no expectation on the link between income diversification and MCBs efficiency. The evidence suggests that the business model matters in influencing MCBs cost efficiency. The coefficients is positive, implying that Italian MCBs would gain from diversifying their business other than intermediation within the income statement (income diversification). This conclusion is robust to every sample of MCBs used in the sensitivity analysis (Table 5). With regards to loan diversification the evidence is mixed. In the five MLM specifications used in Table 4 for the entire sample of MCBs, the estimated parameter is not significant, inducing no interpretation. However, this average effect hides some specificities that the sensitivity analysis helps to capture. Indeed, cost efficiency decreases with loan diversification as far as the sample of southern MCBs is concerned, while the contrary happens for the MCBs operating in the North of Italy. This is a clear signal that location matters. A different impact also exists across efficiency and size distributions. For instance, restricting the regression to the 1181 observations lying in the 3rd and 4th quartiles of efficiency distribution (column 4 of Table 5), the effect of loans diversification becomes positive, while it is negative in the 1st quartile (column 3). Similarly, the impact turns to be positive in the middle of the size distribution (column 7).

Another aspect to be addressed regards the role played by the capital structure. As matter of fact, the financial capital is related to exposure to risk, in a sense that the more indebted a bank the higher the risk of failure, that arises in situations of systemic crisis (Acharya and Viswanathan 2011). In other words, less equity implies higher risk taken and greater leverage, which results in higher borrowing costs. Again, a high level of leverage directly affects funding costs, since paid interests imply less profitability for the bank in the income statement (Berger and Mester 1997). From these arguments, it is reasonable to assume that more leveraged MCBs face high funding costs and then low efficiency scores. In our regressions the capital structure is proxied by the ratio equity/total assets, which ranges from 0 (highly leveraged MCBs) to 1 (financial independent MCBs). The coefficient of the equity/total assets ratio is negative, suggesting that an increased amount of capital, for instance as requirements of regulation, can act as a binding restriction and thus is perceived by MCBs as a cost. It is worth noticing that the negative relationship between equity/total assets and MCBs cost efficiency is robust to any sensitivity check we perform (Table 5)

Turning back to the specific objective of the paper, it is worth discussing the empirics about how the provincial market conditions affect MCBs performance. Results presentation begins with the market concentration, which enters into regressions to gauge the effect of consolidation process observed in Italian banking. This is an issue also addressed, among many others, by Casu and Girardone (2009) Dongili et al. (2008),Fontani and Vitali (2007) and Turati (2008). The uncertainty of the outcome is due to the fact that, on the one hand, the consolidation increases individual size with an expected increase in efficiency levels. On the other hand, high concentration may cause an increase of banks market power and, therefore, a reduction of banks efficiency (Turati 2008). Concentration is measured using the Herfindahl Index and Total Assets (HH2) in each province (cfr § 2.3). From our regressions it emerges that banking concentration is positively related to MCBs efficiency, meaning that MCBs operating in provinces with more concentrated banking markets attain higher cost efficiency. This evidence is robust to every check (Table 5) and consistent with the efficient structure hypothesis (Berger 1995; Goldberg and Rai 1996). Phrased differently, in local concentrated banking markets, each MCB is induced to be more and more efficient, with the result that in provinces with high market concentration there would be a dominance of efficient MCBs. Arguments that increased market concentration leads to efficiency improvements are also provided by Casu and Girardone (2009) and Demirgüç-Kunt and Levine (2001).

Regarding the spatial accessibility to banking services, it is reasonable to argue that banking efficiency in local market can also be affected by the branching that has occurred in Italy over the last 20 years. In more detail, it can be expected that the higher the number of branches the less MCBs efficiency. This is why a large number of branches exerts negative effects of individual efficiency because the operating costs to provide banking services increase. Moreover, local markets with a high number of branches would suffer from over dimensioning which acts against efficiency. However, the sign may be different, as the big-bank participation in small markets can be positive due to the increases in the capital brought by big banks, the expertise brought in risk management and increases in competition (Delis and Papanikolaou 2009; Hannan and Prager 2009). This phenomenon is measured province-by-province with the number of bank branch by square kilometer. Results are in line with the expectation as the estimated parameter of Branch Density is always negative (Tables 4 and 5), indicating that Italian MCBs suffer from the huge branch opening process occurring throughout the country. This evidence might be due to the fact that the presence of many bank branches in local markets forces individual MCB to invest increasing amount of resources for serving more customers, whose expectations is to increase the benefits from loans and deposits at better advantageous conditions than those applied by other bank branches. Other things being fixed, the increased number of bank branches in local markets and the MCBs strategies act against their costs.

Another issue that the study addresses is the effect on efficiency due to demand effects. The hypothesis is that MCBs that operate in markets with a lower density of demand (calculated as deposits per square kilometers) face higher expenses to find customers asking for banking services (Fries and Taci 2005). Thus, the higher the density of demand, the higher will be the banking efficiency levels. These effects are gauged by the demand density expressed as total deposits by square kilometer. From estimations it emerges that MCBs cost efficiency is positively related to demand density (Table 4). This supports the hypothesis according to which MCBs working in provinces with high level of deposits face, *ceteris paribus*, lower costs in mobilizing deposits and making loans. Interesting, the positive link remains positive only in the middle of cost efficiency and size distributions, while the evidence is inconclusive in the tails (Table 5).

In order to gauge the effects of systemic market risk on individual efficiency, in the MLM equation we insert the variable Market Risk, expressed as the nonperforming loans as share of total loans. It is calculated by taking into account the localization of customer in every province. Here, the question to be understood is whether MCBs gain or lose from operating in local markets with poor credit-quality. It is likely that MCBs operating in risky markets are exposed to potential efficiency losses caused by higher costs of screening and monitoring activities. Results differ according to the MLM specification. If the time-effect introduces disturbances in slopes (Model 5 of Table 4), then MCBs cost efficiency will be positively related to the local financial markets riskiness. This finding is driven by banks lying in the upper tail of size distribution, while is robust to efficiency distribution and MCBs location (Table 5). This might be due to the fact that MCBs save costs from the nature of the relationships with their member-customers. These relationships protect MCBs from market riskiness as they are long-dated and based on the use of soft-information.

Finally, the level of local economic development is an important factor of bank performances, because it affects numerous aspects related to the demand and supply of banking services (mainly deposits and loans). To this end, the income per capita is used as measure of local development. It is excepted that provinces with higher income per capita are assumed to have a banking system operating in a mature environment and resulting in more competitive interest rates and profit margins. They can also exert more financial activity. Results are mixed and not robust, given that a significant link has been found only in Models 2 and 3 of Table 4. Contrasting with the expectations, our evidence may be affected by the fact that operating in rich areas implies higher operating and financial costs that MCBs would incur in offering services (Dietsch and Lozano-Vivas 2000).

Table 4
Explaining heterogeneity in cost efficiency of Italian MCBs. Evidence from MLMs with bank and provincial-specific variables over the 2006-2011 period.

		Model 1	Model 2	Model 3	Model 4	Model 5
	Constant	3.183***	3.144***	3.272***	3.172***	3.240***
		(15.99)	(15.71)	(15.28)	(15.74)	(15.63)
Fixed-effects						
	Time		-0.010	-0.0127*	-0.0447***	-0.0435***
			(-1.75)	(-2.20)	(-5.50)	(-5.37)
	MCBs level					
	Size	-0.198***	-0.193***	-0.206***	-0.190***	-0.197***
		(-13.03)	(-12.55)	(-12.41)	(-12.36)	(-12.40)
	Loans Diversification	0.159	0.136	0.124	0.00939	0.0108
		(1.59)	(1.34)	(1.19)	(0.09)	(0.11)
	Income Diversification	3.691***	3.643***	3.680***	3.395***	3.442***
		(33.31)	(32.02)	(32.63)	(29.66)	(30.20)
	Equity/Total Assets	-3.654***	-3.569***	-3.799***	-3.453***	-3.569***
		(-12.96)	(-12.53)	(-13.43)	(-12.27)	(-12.75)
	Province level					
	Market Concentration	0.200**	0.241***	0.229***	0.178***	0.179***
		(5.03)	(5.20)	(5.04)	(3.66)	(3.76)
	Branch Density	-121.84***	-129.9***	-128.4***	-168.0***	-163.8***
		(-4.07)	(-4.31)	(-4.13)	(-4.37)	(-4.30)
	Demand Density	0.002**	0.003***	0.003***	0.004***	0.004***
		(3.24)	(3.61)	(3.50)	(3.91)	(3.85)
	Market Risk	-0.220	0.221	0.745	2.275***	2.258***
		(-0.73)	(0.56)	(1.85)	(4.90)	(4.93)
	Local Econ. Development	-0.018	-0.019**	-0.015*	-0.002	-0.002
		(-2.40)	(-2.59)	(-2.07)	(-0.27)	(-0.29)
Random-Effects						
	Variance					
	Provinces (Intercept)	0.0371	0.0357	0.0373	0.0456	0.0466
	Provinces (Slope)				0.0015	0.0014
	MCBs (Intercept)	0.0435	0.0432	0.0409	0.0430	0.0446
	MCBs (Slope)			0.0014		0.0006
	Time Random Effect	0.0608	0.0686	0.0537	0.0545	0.0518
	Total	0.1414	0.1475	0.1333	0.1446	0.1451
	ICC					
	Provinces	26.23%	24.22%	27.98%	32.57%	33.08%
	MCBs	30.76%	29.29%	31.72%	29.74%	31.19%
	Time	43.01%	46.49%	40.30%	37.68%	35.73%
	R^2	0.3368	0.3053	0.3721	0.3189	0.3169
	R2 level 3	0.2761	0.4049	0.3786	0.2153	0.2006
	R2 level 2	0.1330	0.1385	0.1570	0.1423	0.0980
	R2 level 1	0.0469	0.3286	0.4740	0.4665	0.4927
	LR test (p value)	0.000	0.000	0.000	0.000	0.000
	Log restricted-lik	-435.01	-437.77	-416.01	-372.17	-365.54
	AIC	896.03	903.53	862.03	774.35	763.07
	Number of Groups					
	Provinces	66	66	66	66	66
	MBCs	414	414	414	414	414
Number of observations		2334	2334	2334	2334	2334

Table 5

Explaining heterogeneity in cost efficiency of Italian MCBs. Evidence from MLMs with bank and provincial-specific variables. A sensitivity analysis (2006-2011)

	Banks Location		Cost Efficiency Distribution			Size Distribution (by quartile)		
	North	South	1st	2nd and 3rd	4th	1st	2nd and 3rd	4th
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Constant	2.512***	4.442***	4.634***	1.419**	2.012***	0.699**	1.917***	2.852***
	(10.95)	(8.95)	(5.84)	(2.90)	(3.30)	(3.13)	(16.71)	(15.78)
Fixed-effects								
Time	-0.0298***	-0.0891***	-0.0658***	-0.0369***	-0.00952	0.0115	-0.00152	-0.0416***
	(-3.39)	(-5.26)	(-3.93)	(-3.61)	(-0.71)	(1.13)	(-0.33)	(-5.84)
MCBs level								
Size	-0.159***	-0.274***	-0.294***	-0.0733	-0.120**	-0.0189	-0.059***	-0.083***
	(-9.40)	(-6.71)	(-4.12)	(-1.86)	(-2.62)	(-1.13)	(-6.88)	(-5.69)
Loans Diversification	0.330**	-0.704***	-0.487*	0.552***	0.247	0.158	0.158*	-0.101
	(2.70)	(-3.59)	(-2.47)	(3.75)	(1.12)	(1.20)	(2.36)	(-1.09)
Income Diversification	3.438***	3.528***	3.678***	3.743***	3.282***	1.589***	1.058***	1.049***
	(24.97)	(16.27)	(14.71)	(22.9)	(17.16)	(8.05)	(10.18)	(7.83)
Equity/Total Assets	-1.661***	-4.376***	-4.067***	-1.02	0.345	-1.492***	-1.061***	0.151
	(-3.38)	(-10.92)	(-8.48)	(-0.62)	(0.16)	(-4.70)	(-4.48)	(0.57)
Province level								
Market Concentration	0.178***	0.302*	0.329*	0.271***	-0.049	-0.001	0.054	0.196*
	(3.57)	(1.97)	(2.33)	(4.58)	(-0.73)	(-0.02)	(1.45)	(2.40)
Branch Density	-114.8**	-57.66	-137.4	-202.8***	-48.32	25.86	-50.10**	-72.22*
	(-3.04)	(-0.25)	(-1.31)	(-3.83)	(-1.05)	-0.9	(-2.65)	(-2.13)
Demand Density	0.003**	0.017**	0.004	0.005**	0.001	-0.001	0.001*	0.002
	(2.66)	(2.82)	(1.08)	(3.16)	(1.14)	(-0.78)	(2.35)	(1.86)
Market Risk	2.697***	1.630*	1.917*	1.696**	2.096*	0.685	0.368	0.769*
	(4.13)	(2.12)	(2.20)	(2.80)	(2.09)	(1.12)	(1.22)	(1.98)
Local Econ. Development	0.009	-0.014	-0.031	0.011	0.001	0.008	0.002	-0.019*
	(1.00)	(-1.09)	(-1.70)	(1.09)	(0.01)	(0.69)	(0.38)	(-1.96)
Random-Effects								
Variance								
Provinces (Intercept)	0.0320	0.0457	0.0025	0.0035	0.0040	0.0456	0.0506	0.0657
Provinces (Slope)	0.0010	0.0015	0.0003	0.00006	0.00006	0.0010	0.0014	0.0020
MCBs (Intercept)	0.0370	0.0472	0.0130	0.0051	0.0065	0.0323	0.0479	0.0246
MCBs (Slope)	0.0003	0.0015	6.21E-23	0.00010	2.95E-25	0.0009	0.0005	0.0006
Time Random Effect	0.0520	0.0478	0.0400	0.0189	0.0249	0.0871	0.0383	0.0331
Total	0.1223	0.1437	0.0558	0.0277	0.0355	0.1669	0.1387	0.1260
ICC								
Provinces	26.98%	32.85%	5.02%	12.86%	11.44%	27.94%	37.49%	53.73%
MCBs	30.50%	33.89%	23.30%	18.79%	18.33%	19.87%	34.90%	19.99%
Time	42.52%	33.26%	71.68%	68.35%	70.22%	52.19%	27.61%	26.27%
LR test (p value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Log restricted-lik	-239.50	-95.34	15.81	472.82	158.78	-218.13	-78.95	-7.11
AIC	510.99	222.69	0.39	-913.63	-285.56	468.25	189.90	46.22
Number of Groups								
Provinces	42	24	57	65	58	46	60	42
MBCs	312	102	222	380	232	122	241	119
Number of observations	1784	550	584	1166	584	557	1181	596

5. Concluding remarks

This paper investigates the efficiency of MCBs over the 2006-2011 period by using the multilevel approach for longitudinal data. The underlying idea is that MCBs efficiency is the result of individual behavior and of local market-specific conditions. While the existing studies on this topic use single-equation model, we combine contextual and micro links within the multilevel model, which properly handle the embeddedness of MCBs in geographically narrowed markets. The preferred model is an equation where intercepts randomly vary across MCBs and provinces and time is modeled as a source of randomness of intercepts and slopes. As the main research-question is about the role of local markets, the descriptive analysis of data allows to reveal that several significant differences exist at provincial level: despite the reforms, the Italian banking market is still highly heterogeneous, with marked differences between the North and the South of the country. Importantly, the main results of the study are shown to be robust with regard to model specification and across different samples of banks.

The study yields two main results.

The first evidence regards the role of localization in different provinces. From this perspective, several points stand out. Above all, we find that the heterogeneity in efficiency can, to a large extent, be explained by unobserved province-specific factors. For instance, in the empty model, provinces explain about one third of efficiency heterogeneity, while this proportion is one fifth in the most extended model (the one with observables at individual and contextual levels and with time-randomness in intercepts and slopes). Furthermore, the study emphasizes the positive relationship between efficiency and market concentration. Other robust insights come from the demand density and the branch density, which positively and negatively affect cost efficiency. The evidence from credit quality indicates that MCBs efficiency appears not to be related to the riskiness of local banking markets. When trying to summarize the effects of local market conditions, the analysis induces an indirect assessment of deregulation, although the study is not centered on evaluation. In this respect, the finding that high market concentration is positively linked to MCBs efficiency could be considered as an implication of reforms carried out over the last 15 years. In this sense a virtuous-circle seems to be at work: market concentration in the periphery makes MCBs in those markets be more efficient and then viable. This is in line with the intentions of regulators, as the scope to maintain market efficiency is an expected result of market consolidation. At the same time, MCBs viability preserves the small market to be served. However, the negative effect of branching on MCBs efficiency acts against the full effectiveness of reforms, as the impressive branch opening is seen as a threat for efficiency and thus MCBs survival.

Secondly, heterogeneity in cost efficiency is affected by bank-specific factors. For instance, in the empty model, the proportion of MCBs efficiency variability brought about by the bank-level of our hierarchy is high, ranging from one quarter in the empty model to one third in the full-mixed model. While these results imply that the unobserved heterogeneity in MCB-behavior is an important source of heterogeneity in efficiency, they should be looked at in greater depth. In this respect, regressions incorporate the effect of a set of bank-specific variables relating to size, diversification and capital structure. The lessons we have learnt are twofold. On one hand, looking at the impact on efficiency exerted by each factor, we find MCBs efficiency always increases with income diversification and when small-banks are financial dependent on external finance, while it is negatively related with size, although the marginal effect is rather small. On the other hand, we evaluate the capacity of the above bank-level variables to explain the total efficiency variability. We show that the bank-specific variables explain, as a whole, nine per cent of first-level efficiency variance, implying that much of efficiency heterogeneity at individual basis is still unexplained. Something other than size, diversification and capital structure influences heterogeneity in MCBs efficiency. This leaves room for further research with the aim of refining the measurement issues relating to other bank-level aspects, such as management competence and organizational practices. It would be interesting to analyze these issues in greater depth so as to minimize the "sizable" and "unobservable" black box of MCBs behavior. On one hand, the paper suggests that MCBs-based reforms could be highly advantageous in terms of efficiency gains, as they would act within the level that this study demonstrates, as expected, to be an important dimension in explaining the efficiency heterogeneity. Limiting the discussion to the organizational efficiency nexus, it seems that the policy making might be better oriented to stimulate organizational changes of MCBs as a whole, preserving, however, the specific nature of being small banks. Hence, any reform ought to be oriented to assure customers in the periphery to be served. In this, a reform allowing MCBs to use soft information and lean-relationship could be a good option, as it would guarantee some advantages to MBCs over big banks.

Appendix Table A Territorial breakdown of MCBs, by province(2006-2011)

Provinces	Number of MCBs observations	Provinces	Number of MCBs observations
Agrigento	20	Macerata	18
Ancona	42	Mantova	14
Arezzo	12	Matera	2
Ascoli Piceno	22	Messina	12
Avellino	18	Milan	60
Bari	76	Padova	48
Benevento	11	Palermo	32
Bergamo	54	Perugia	24
Bologna	35	Pesaro-Urbino	36
Bolzano	292	Pisa	13
Brescia	63	Pistoia	42
Brindisi	12	Pordenone	12
Caltanissetta	36	Potenza	23
Campobasso	13	Ravenna	12
Caserta	12	Reggio Emilia	18
Catania	12	Rimini	22
Catanzaro	23	Rome	67
Chieti	12	Rovigo	24
Como	18	Salerno	81
Cosenza	34	Siena	31
Cremona	28	Syracuse	19
Crotone	9	Taranto	29
Cuneo	46	Teramo	22
Florence	42	Trapani	18
Forlì-Cesena	36	Trento	275
Frosinone	17	Treviso	36
Gorizia	25	Udine	48
Grosseto	24	Venezia	24
Latina	22	Verona	42
Lecce	12	Vibo Valentia	12
Lecco	12	Vicenza	60
Livorno	15	Viterbo	22
Lodi	18		
Lucca	13	Italy	2334

Source: our elaborations on data from ABI and Bank of Italy.

References

Acharya V.V., Viswanathan S. (2011) "Leverage, Moral Hazard, and Liquidity", *Journal of Finance*, 46, 99-138

Aiello F., Pupo V.,Ricotta F. (2014) "Explaining Total Factor Productivity at Firm Level in Italy: Does Location Matter?", *Spatial Economic Analysis*, 9 (1),1-20

Aiello F., Pupo V.,Ricotta F. (2015) "Firm Heterogeneity in TFP, sectoral innovation and location. Evidence from Italy", *International Review of Applied Economics* 29(5),546-571

Aigner D., Lovell C.A.K., Schmidt P. (1977) "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, 6, 21-37

Alessandrini P., Presbitero A.F., Zazzaro A. (2009) "Banks, Distances and Firms' Financing Constraints", *Review of Finance*, 13, 2, 261-307

Andries A.M. (2011) "The Determinants of Bank Efficiency and Productivity Growth in the Central and Eastern European Banking Systems", *Eastern European Economics*, 49, 6, 38-59

Baele L., De Jonghe O., Vennet R.V. (2007) "Does the Stock Market Value Bank Diversification?", *Journal of Banking and Finance*, 31, 7, 1999-2023

Baldwin R.E., Okubo T. (2006) "Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting", *Journal of Economic Geography*, 6,323-346

Barra C., Destefanis S., Lubrano Lavadera G. (2014), The Crisis and the Efficiency of Italian Cooperative Banks, Paper Presented at the 55th Annual Meeting of the Italian Economic Society, 23-25 October 2014, Trento (Italy)

Battaglia F., Farina V., Fiordelisi F., Ricci O. (2010) "The Efficiency of Cooperative Banks: the Impact of Environmental Economic Conditions", *Applied Financial Economics*, 20, 17, 1363-1376

Battese G.E., Coelli T.J. (1995) "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", *Empirical Economics*, 20, 2, 325-332

Berger A.N. (1995) "The Profit-Structure Relationship in Banking-Tests of Market-Power and Efficient-Structure Hypotheses", *Journal of Money, Credit and Banking,* 27, 2, 404-431

Berger A.N., Mester L.J. (1997) "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?", *Journal of Banking and Finance*, 21, 7, 895-947

Beugelsdijk S. (2007) "The regional environment and a firm's innovative performance: a plea for a multilevel interactionist approach", *Economic Geography*, 83,181-199

Bickel R. (2007) *Multilevel Analysis for Applied Research, It's Just Regression*!, The Guilford Press, New York

Bos J.W.B., Kool C.J.M. (2006) "Bank Efficiency: the Role of Bank Strategy and Local Market Conditions", *Journal of Banking and Finance*, 30, 1953-74

Carbò Valverde S., Humphrey D.B., Rodriguez F.R. (2003) "Deregulation, Bank Competition and Regional Growth", *Regional Studies*, 37, 3, 227-237

Casu B., Molyneux P. (2003) "A Comparative Study of Efficiency in European Banking", *Applied Economics*, 35, 17, 1865-1876

Coccorese P. (2009) "Market Power in Local Banking Monopolies", *Journal of Banking and Finance, 33, 1196-2120*

Delis M., Papanikolaou N. (2009) "Determinants of Bank Efficiency: Evidence from a Semi-parametric Methodology", *MPRA*, Working Paper N. 13893

Demirgüç-Kunt A., Levine R. (2001) *Bank-based and Market-based Financial Systems: Cross-Country Comparisons*. In Demirgüç-Kunt A., Levine R. (Eds.): Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets, and Development. MIT Press, Cambridge

Dietsch M., Lozano-Vivas A. (2000) “How the Environment Determines Banking Efficiency: a Comparison between French and Spanish Industries”, *Journal of Banking and Finance*, 24, 985-1004

Dongili P., Rossi S.P.S., Zago A. (2008) “Efficienza e competitività delle banche italiane: un confronto con alcuni sistemi bancari europei”. In Cella G., Zago A. (Eds.): *Competitività ed efficienza dell’economia italiana: fattori sistemici e valutazioni quantitative*. Il Mulino: Bologna

Dosi G., Nelson R. (2010) “Technical Change and Industrial Dynamics as Evolutionary Processes.” In *Handbook of the Economics of Innovation*, Hall, and N. Rosenberg (Eds). London: North Holland

Drake L. (2001) “Efficiency and Productivity Change in UK Banking”, *Applied Financial Economics*, 11, 5, 557-571

Fazio G.,Piacentino D. (2010) “A Spatial Multilevel Analysis of Italian SMEs' Productivity” *Spatial Economic Analysis*, 5(3), 299-316

Fiordelisi F., Mare D.S. (2013) “Probability of Default and Efficiency in Cooperative Banking”, *Journal of International Financial Markets, Institutions and Money*, 26, 30-45

Fontani A., Vitali L. (2007) “L’efficienza di costo dei gruppi bancari italiani: un’analisi mediante frontiera stocastica”, *Department of Economics and Business*, Luiss, Rome

Fries S., Taci A., (2005) “Cost Efficiency of Banks in Transition: Evidence from 289 Banks in 15 Post-Communist Countries”, *Journal of Banking and Finance*, 29, 1, 55-81

Giannola A. (2009) *Bank Mergers and Credit Allocation Among Italian Regions.* In Silipo D.B. (Ed.): The Banks and the Italian Economy. Springer Phisyca-Verlag: Berlin Heidelberg

Girardone C., Molyneux P., Gardener E.P.M. (2004) “Analysing the Determinants of Bank Efficiency: the Case of Italian Banks”, *Applied Economics*, 36, 3, 215-227

Goldberg L.G., Rai A. (1996) “The Structure-Performance relationship for European Banking”, *Journal of Banking and Finance*, 20, 4, 745-771

Goldstein H. 2003. *Multilevel Statistical Models*, London, Arnold

Gutiérrez E. (2008) “The Reform of Italian Cooperative Banks: Discussion and Proposals”, *IMF WP/08/74*, IMF Washington

Heck R.H., Thomas S.L. (2000) *An introduction to multilevel modelling techniques*. Lawrence Erlbaum Associates, Mahwah, New Jersey

Hox J.J. (2002) *Multilevel analysis: techniques and applications*. Lawrence Erlbaum Associates, Mahwah, New Jersey

Hannan T.H., Prager R.A. (2009) “The Profitability of Small Single-Market Banks in an era of Multi-Market Banking”, *Journal of Banking and Finance*, 33, 2, 263-271

Hughes J.P., Mester I.J. (2008) “Efficiency in Banking: theory, practice and evidence”, *WP 08/01*, Federal Reserve Bank of Philadelfia, Philadelfia, USA

Kwok O.-M., Underhill A.T., Berry J.W., Luo W., Elliott T.R., Yoon M. (2008) “Analyzing Longitudinal Data with Multilevel Models: an Example with Individuals Living with Lower Extremity Intra-articular Fractures”, *Rehabilitation Psychology*, 53, 3, 370-386

Little R.J.A., Rubin D.B. (2002) *Statistical Analysis with Missing Data*, 2nd ed. Hoboken, NJ: Wiley

Luke A. (2014) *Multilevel Modelling*. London: Sage

Maddala G.S. (1991) “A Perspective on the Use of Limited-Dependent and Qualitative Variables Models in Accounting Research”, *Accounting Review*, 66, 788-807

Mass J.M, Hox J.J. (2004) ”Robustness issues in multilevel regression analysis. Statistica Neederlandica 58: 127-137

Maudos J., Pastor J.M., Perez F., Quesada J. (2002) “Cost and Profit Efficiency in European Banks”, *Journal of International Financial Markets, Institutions and Money*, 12, 1, 33-58

McDonald J. (2009) “Using Least Squares and Tobit in Second Stage DEA Efficiency Analyses”, *European Journal of Operational Research*, 197, 792-798

Meeusen D., van de Broek J. (1997) “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error”, *International Economic Review*, 18, 2, 435-444

Messori M., Tamburini R., Zazzaro A. (2003) *Il sistema bancario italiano*, Carocci Editore: Rome

Pilloff S.J. (1996) “Performance Changes and Shareholder Wealth Creation: Associated with Mergers of Publicly Traded Banking Institutions, *Journal of Money, Credit and Banking*, 28, 3, 59-78

Rabe-Hesketh S., Skrondal A. (2008) Multilevel and Longitudinal Modeling Using Stata. Stata Press, 2nd edition

Raspe O.,F. van Oort F. (2011) “Firm Heterogeneity, Productivity, and Spatially Bounded Knowledge Externalities”, *Socio-Economic Planning Sciences* 45: 38-47

Richter T. (2006) “What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models”, *Discourse Processes*, 41(3): 221-250

Romer P. 1986. “Increasing returns and long-run growth.” *Journal of Political Economy* 94: 1002-1037

Royston P. (2007) “Multiple Imputation of Missing Values: Further Update of Ice, with an Emphasis on Interval Censoring”, *Stata Journal*, 7, 445-464

Royston P. (2009) “Multiple Imputation of Missing Values: Further Update of Ice, with an Emphasis on Categorical Variables”, *Stata Journa*l, 9, 466-477

Silipo D.B. (2009) *The Banks and the Italian Economy*. Springer Phisyca-Verlag: Berlin Heidelberg

Snijders T.A.B.,Bosker R.J. (2008) “Diagnostic checks for multilevel models” in Handbook of multilevel Analysis, Jan de Leeuw (Eds.), New York: Springer

Srholec, M. 2010. “Multilevel Approach to Geography of Innovation.” *Regional Studies* 44:1207-1220.

Srholec M. (2015) “Understanding the Diversity of Cooperation on Innovation across Countries.” *Economics of Innovation and New Technology* 24(1-2): 159-182

Turati G. (2008) “La valutazione del grado di concorrenza nell’industria bancaria negli anni Novanta”. In Cella G., Zago A. (Eds.): *Competitività ed efficienza dell’economia italiana: fattori sistemici e valutazioni quantitative*, Il Mulino: Bologna

Vallascas F., Keasey K. (2012) “Bank Resilience to Systemic Shocks and the Stability of Banking Systems: Small is Beautiful”, *Journal of International Money and Finance*, 31, 1745-1776

van Oort, F.G., Burger M.J., Knoben J., Raspe O. (2012) “Multilevel Approaches and the Firms Agglomeration Ambiguity in Economic Growth Studies”, *Journal of Economic Surveys*, 26: 468-491

Wagner W. (2010) “Diversification at Financial Institutions and Systemic Crises”, *Journal of Financial Intermediation*, 19, 373-386