COMPETITIVE ISSUES IN BANKING: THEORETICAL APPROACHES OVERVIEW

IRYNA ALOSHYNA
National Taras Shevchenko University of Kyiv, 02032, Ukraine
ira_galchenko@email.ua

Abstract
Theoretical concepts of bank competitiveness are studied and research approaches to the banking efficiency measurement are overviewed. Empirical models of competition in banking are considered for identifying the key factors of bank competitiveness.

Key words: bank efficiency, bank competitiveness, bank productivity, competition in banking, competitive market.

JEL Classification: G21

I. INTRODUCTION

The development of global financial markets has acted to alter the competitive environment of financial institutions (Hirtle, 1991, p.38). Thus internationally active financial institutions (banks) should respond to opportunities and challenges of an increasingly competitive global market environment due to increased access of foreign competitors to domestic financial markets and the expanding availability of traditional banking services from non-traditional sources. In such circumstances the intense competition have changed the nature of market for the existing banking products and discovered the new ones where financial institutions should compete on national and international level. Furthermore increasing integration of banks into financial markets allows banks to shift their traditional risks to the markets which influence on financial system stability in general (Boot and Thakor, 2009, p.21).

The aim of the research is to study the concepts of bank competitiveness and the models of its measurement. The research objectives are:

- to study competitive issues according to bank productivity studies in order define simple production function and its efficiency frontiers;
- to consider the empirical models of banking competition in order to identify major factors of bank competitiveness.

II. COMPETITIVE ISSUES ACCORDING TO BANK PRODUCTIVITY STUDIES

Competition is a rivalry between individuals of companies which are interested in a reaching its goals. From the social welfare point of view perfect competition may not be optimal for all industries. Competition in the financial system is limited considering the welfare of the financial system and to avoid large rescue costs,. Banks are somewhat more protected than other industries to ensure that their profitability is secured so that they have an incentive to take risks (Yakoi-Arai and Yoshino, 2006, p.8).

That’s why the standard paradigm of perfect competition is not considered appropriate for banking as the free entry and exit of banks from the market is not possible. Banking sector is prone to becoming oligopolic because of prudential regulation such as licensing and capital requirements. Thus competition in banking has some specific features which distinguish it from the competition in other industries:

- banking is leaned toward oligopoly model which is more appropriate from the effectiveness point of view;
- the subjects of competition are not only banks, but also other financial and non-financial organizations;
- competitive environment is a variety of economy sectors in which banks could be both a seller or a buyer;
- competition is exercised in different forms: the intense intrasectoral competition connected to the diversification of banking products and low intersectoral competition because of absent of non-banking competitiveness substitutes of banking products;
intense competition limits ability to use price methods and arises the topicality of quality control.

The topic of competition in banking is less straightforward because of the intangible nature of banking ranging from non-price features associated with virtually all bank services to the maturity structure of bank assets and liabilities (Heffernan, 2005, p.473). The measurement of output of services produced by banks is specially problematic because of difficulty to account for its quality. That is why there are two different approaches to measure bank output (Heffernan, 2005, pp.474-477; Bikker, 2005, p.115):

1) The production approach: banks are treated as firms which use capital and labour to produce different categories of deposit and loan accounts, its outputs are measured by the number of these accounts or the number of transactions per account. According to this bank output is treated as a flow, that is, the amount of output produced per unit of time. Total costs are all operating costs (the cost of factor input such as labour and capital) used to produce these outputs. But such approach has several problems: how to weight each bank service in the computation of output; the method ignores interest costs and inflation rates; data from banks using different accounting systems in different countries may not be comparable;

2) The intermediation approach: banks are not producers of loan and deposit, they are intermediators, its outputs are measured by the value of loans and investments. According to this bank output is treated as a stock, showing the given amount of output at one point in time. Total costs are all operating costs used to produce these outputs plus interest costs.

Thereby the intermediation approach because of the fewer data problems is used in most bank productivity studies. Such studies introduce banks as rational economic agents (Bikker, 2005, pp.116-120):

- banks act rationally – operate in such a way that it pursues its own goals in what is the best conceivable, optimal way;
- banks are agents interacting with other agents such as consumers and governments;
- banks are economic agents – their goals are defined in economic terms. Thus banks try to be productive and efficient.

Figure 1 – Bank productivity

In order to explain productivity and efficiency concept we use a simple production function $y=f(x)$, where output $y$ is produced using input $x$, the production function is twice continuously differentiable. Productivity is defined here as the rate at which output changes as all inputs are varied (if as a result of doubling each of the tree factor inputs the bank is able to more than double its loan portfolio). Economies of scope exist if the joint production costs of producing two or more outputs are lower than if the products are produces separately.

Maximum output

Figure 2 – Bank efficiency

Efficiency is defined as productivity – the maximum output for a given level of time – the difference between observed and optimal input/output mixes: $(y/x)/(y^*/x^*)=y/y^*$, where $y$ is the vector of outputs $y_1$ and $y_2$ (Fig.2). Such type of efficiency is referred to X-efficiency that results from the position of bank within the feasible production set and relative to the production frontier. X-efficiency can be measured in terms of cost and profit (Heffernan, 2005, pp.477-478):
Cost X-efficiency: objective of bank to be on the frontier, not above it, not possible to be on points below the curve

Profit X-efficiency: objective of bank to move the maximum profit point on the profit frontier curve

Figure 3 – Cost and Profit X-efficiency

Thus profit-maximizing banks will want to be on the efficiency frontier but can go further unless there is technological change which moves it downwards. However in markets with some degree of monopoly power (due to entry barriers, scale economies or regulation) some banks are located somewhere inside the X-efficiency frontier.

Cost-minimizing banks will want to be on the efficiency frontier, which they could reach due to scale and scope economies that provide measures of the extent to which unit costs could be lowered by offering the total volume of production or range of products. Lower X-efficiency means reducing costs through improved management or greater employee productivity which moves bank closer to the most efficient way of harnessing a given set of resources (Heffernan, 2005, p. 478).

Thus we have defined the simple production function of banks and the frontiers of its efficiency. In banking it is very important to identify special problems which arise from the equivocation in the production function of banks (Bikker, p.121). Such equivocation is the reason of efficiency measurement modelling based on cost and/or profit function which could explain the result of banks.

According to Heffernan, there are two different ways of testing for X-efficiencies (Heffernan, 2005, pp.478-482):

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Non-parametric approach</th>
<th>Parametric approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodology</td>
<td>Data envelop analysis – is not based on any explicit model of the frontier because accounting profit measures are difficult to compute.</td>
<td>Stochastic frontier approach – uses a translog cost function.</td>
</tr>
<tr>
<td>Measuring formula</td>
<td>Compares observed outputs (Y_{jp}) and inputs (X_{jp}) of several organizations, the relatively more efficient bank can be compared against the relatively less efficient by identifying a ‘best practice’. To do this, maximize the following: [E_p = \sum u_j y_j / \sum u_j x_j] subject to (E_p \leq 1) for all (p), where (p) presents several banks and weights (u_j &gt; 0). Thus each bank will have a derived rating of (E) - a measure of relative efficiency, the closer (E) to 1, the higher the relative efficiency.</td>
<td>Involves estimating a cost (or profit) function for a sector: [TC = TC(q, p, y, z, \mu, \epsilon)] (2), where (TC) – variable total costs, (q) – a vector of quantities of variable outputs, (p) – factor inputs, (y) – other variables (environmental or market) which effect output, (z) – quantities of fixed inputs of outputs which effect variable costs, (\mu) – error term which can arise because the bank fails to react optimally to the vector of input prices or from employing too many of the input to produce (q). Using natural logs (ln) on both sides of the equation gives (3): [\ln TC = f(q, p, y, z) + \ln \mu + \ln \epsilon] Profit efficiency measures show how close a bank is to producing the maximum profit possible given input prices, output prices and other variable (4): [\ln(\pi + \theta) = F(q, x, y, z) + \ln \mu + \ln \epsilon] where (\pi) – variable profits, (\theta) – constant to ensure the natural log is of positive number, (x) – vector of prices of the variable outputs, (\mu) – inefficiency that reduces profits.</td>
</tr>
<tr>
<td>Usage in banking</td>
<td>Via linear programming model it measures the efficiency frontier, the most efficient bank is on its frontier and the most inefficient – outside.</td>
<td>A bank is inefficient if its costs exceed those of the most efficient bank using the same input-output combination</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
<td>Can very over time and all outputs and inputs are handled simultaneously which produces true frontier of efficiency</td>
<td>Allows a more breakdown of the constituents of X-efficiency, namely technical inefficiency which arises from factor inputs being over-used and allocative inefficiency due to lax management or expense behaviour.</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>Random errors from measurement problems: inaccurate data, random effects on outputs or inputs, specification errors, efficiency scores are not independent of market structure.</td>
<td></td>
</tr>
</tbody>
</table>

Thus, thanks to this table we can discover the methodology of parametric and non-parametric approaches of testing X-efficiency in banking and its formulas can help us to understand the following models of competition in banking sector.

**III. EMPIRICAL MODELS OF COMPETITION IN BANKING**

Empirical models of competition in banking are used to assess how competitive the banking sector is and help to identify factors influencing competitive structure. This empirical research can be subdivided into (Heffernan, 2005, pp.494-514; Bikker, 2005, pp.105-109; Bikker, 2007, pp.4-12, Degryse and Ongena, 2004, pp.5-22):

**Bank as a traditional industrial organization:**

1) **Structure-Conduct-Performance Model** (Bain 1956; Berger and Hannan 1989, 2004; Altunbas and Gardener 1997; Punt and Rooij 2003), according to which a change in the market structure (the number of banks, their level of concentration, the opportunities for market entry) affects the way banks behave (mutual pricing agreements) and perform (monopoly rents then lead to excessive profitability). The more concentrated the market, the more market power banks have which means they can be inefficient without being forced out of the market. This approach assumes a link between:

\[
P = f (\text{CONC}, \text{MS}, D, C, x) \quad (5),
\]

where \(P\) – performance, \(\text{CONC}\) – marker structure (market concentration), \(\text{MS}\) – market share (more efficient banks should have bigger market share), \(D\) – demand on the market, \(C\) – variables which reflect cost difference, \(x\) – different control variables.

Such model supports the government intervention in order to limit monopoly power and decrease market concentration.

2) **The Relatively Efficiency Model** (Demsetz 1973; Peltzmann 1977; Berger 1995), according to which some banks earn supernormal profits because they are more efficient than others which is reflected in greater output. Such model, as the previous one, says that the more concentrated market generates greater profits but such profits are correlated with efficiency:

\[
\text{Efficiency} \rightarrow \text{Conduct (higher output and/or lower prices)} \rightarrow \text{Market share} \rightarrow \text{Performance (higher profits)}
\]

According to this model, prices and concentration, the opposite SCP, are inversely related. Under the relative efficiency hypothesis, causation runs from greater efficiency, lower prices and higher concentration/market share. Such approach can be linked to the: A). X-efficiency hypothesis – some banks have superior management or production technology which makes them relatively more cost efficient with lower costs, they are able to offer lower prices, gain market share and earn more profit; B). Scale economies hypothesis – banks produce at a low unit costs, lower prices and higher profits per unit of output.
Such model doesn’t support the government intervention into banking but mergers and acquisitions of banks should be encouraged in case when it increase relative efficiency and should be prevented when it increase market concentration.

3). Economies of Scale and Scope Model (Allen and Rai 1996), according to which banks produce the optimal output mix both in terms of size and composition:

$$\ln(TC_i) = f(y_i, p_i) + \varepsilon_i$$ (6),

where $TC$ – total costs, $y$ – outputs, $p$ – input prices of bank $i$ in time $t$.

However such models are being criticized because in there is one-way causality between market structure and performance and is not considered market behaviour and the impact of bank performance on market structure. New Empirical Industrial Organization (NEIO) circumvents this problem and does not try to infer the degree of competition from “indirect proxies”, it aims to infer firms’ conduct directly employing a variety of alternative methodologies with sometimes substantially different data requirements.

**Bank as a new empirical industrial organization:**

1) Contestable Banking Markets Model (Panzar and Rosse 1987; Bikker and Groeneveld 2000; Bikker and Haaf 2002; Bikker 2004), according to which a contestable market is one in which firms are vulnerable to ‘hit and run’ entry and exit and behave as though they are price takers. This type of entry is possible if the market is one where customers can switch suppliers faster than the suppliers can reprice, if incumbents and newcomers have access to similar technology and factor prices and there are no sunk or irrecoverable costs. Thus the players on such contestable market could gain an appropriate market share offering lower prices nevertheless the number of players on that.

To test banking market whether it is contestable in such model it is used the Panzar-Rosse Statistics – the measurement technique of market power through investigating the extent to which changes in factor input prices (labour per unit, real estate per unit, interest costs to total deposits) are reflected in equilibrium industry or bank-specific revenues. To do this consider the following equation of marginal costs and marginal revenue:

$$\ln MC = \alpha_0 + \alpha_1 \ln OUT + \sum_{j=1}^{m} \beta_j \ln FIP_j + \sum_{k=1}^{n} \gamma_k \ln EX_{cos_k}$$ (7),

where $MC$ – marginal costs, $OUT$ – bank output, $FIP$ - factors prices, $EXcost$ – other exogenous factors of cost function.

$$\ln MR = \delta_0 + \delta_1 \ln OUT + \sum_{k=1}^{n} \xi_k \ln EX_{rev_k}$$ (8),


To maximize the revenue marginal costs should be equal to marginal revenue, thus we have the following:

$$\ln OUT = (\alpha_0 - \delta_0) + \sum_{j=1}^{m} \beta_j \ln FIP_j + \sum_{k=1}^{n} \gamma_k \ln EX_{cos_k} - \sum_{k=1}^{n} \xi_k \ln EX_{rev_k} + \rho \ln EX_{PCE}$$. (9),

Based on such theory we have the following formula:

$$\ln H = \alpha + \beta \ln AFR + \gamma \ln PPE + \delta \ln PCE + \sum_{j=1}^{m} \eta_j \ln BSF_j + \varepsilon \ln OI + error$$ (10),

where $AFR$ – the ratio of annual interest expenses to total funds, $PPE$ – ratio of personnel expenses to total balance sheet, $BSF$ – other bank-specific exogenous factors (ratio of other incomes to total assets).

According to this market power is measured as the level of changes of factor prices are reflected on the equilibrium of bank revenue $R_i$. Using the formula (10), The Panzar-Rosse Model tests the H-Statistics hypothesis (the sum of elasticities of the scaled total interest revenue of the banks with respect to their factor input prices) which helps to estimate quantitatively competitive nature of banking markets and their market power:

$$H = \sum \frac{\partial R_i}{\partial \omega R_i}$$ (11).

Perfect competition implies an H-statistic equal to one. A monopoly situation yields an H-statistic that can be negative or zero. Monopolistic competition yields values of H in between zero and one.

Such model doesn’t need government intervention because banks offer their products at marginal costs maximizing customers’ surplus. But it has some restrictions connected with constant improving in technologies which shorten the period between the entry of new player and his decision to exit of the market.

2) Generalised Linear Pricing Model (Heffernan 2002), according to which it is outputted the generalized pricing equation that can be applied to the key retail banking products:

A) Deposits:

$$Rl_i = \alpha_0 + \sum_j \beta_j Libor_{t-j} + \gamma_j + \delta_j D_t + \varphi_i + \varepsilon_i$$ (12),
where \( Rl_i \) - gross deposit rate paid by bank \( i \) at time \( t \), \( Libor_{i-j} \) -\( 0,1,2,3 \) monthly lags used on Libor, the London interbank offer rate, \( n \) - number of banks offering the product, \( t \) - time trend, \( \alpha_i \) - dummy variable for each bank \( i \), unity for bank \( i \).

**B) Mortgages and loans:**

\[
Rl_i = \alpha_0 + \sum_j \beta_j Libor_{i-j} + \gamma_i + \delta_i \epsilon_i + \xi_i + \epsilon_a (13),
\]

where \( \epsilon_a \) - loan or mortgage percentage rate changed by bank \( i \) at time \( t \).

**C) Credit cards:**

\[
Rl_i = \alpha_0 + \sum_j \beta_j Libor_{i-j} + \gamma_i + \delta_i \epsilon_i + \xi_i + \eta_i f_{i} + \epsilon_a (14),
\]

where \( f_{i} \) - fee for credit cards charged by bank \( i \) at time \( t \).

To test for the degree of competition in the banking market is required a benchmark for a perfectly competitive rate against which deposit and loan rates can be compared – the London Interbank Offered Rate, is the rate banks quote each other for overnight deposits and loans. It presents opportunity cost of all of a bank’s assets and the marginal costs of all liabilities. It is an international rate to which all banks have access and therefore is representative of perfectly competitive rate.

**3) Conjectural-Variations Model** (Iwata 1974; Bresnahan 1982; Lau 1982; Shaffer 1993), according to which a bank when choosing its output takes into account the “reaction” of rival banks. The equilibrium oligopoly price is then characterized by the following first order condition:

\[
P(\alpha, \beta) + \lambda \delta P(Q,Y;\alpha) = C(Q, Z; \beta) (15),
\]

where \( P \) - market equilibrium price, \( P(Q,Y;\alpha) \) - market inverse demand function, \( Q \) - market level quantity, \( C(Q, Z; \beta) \) - market marginal cost, \( D \) and \( E \) - vectors of unknown parameters associated with demand and costs respectively, \( Y \) and \( Z \) - vectors of variables that affect demand and costs respectively, \( \lambda \) - conjectural elasticity of total bank industry output to variation of bank \( i \) output (Lerner index).

\[
\lambda = \eta(P) \left[ \frac{\partial P}{\partial MC} \right] (16),
\]

where \( \eta(P) \) , price elasticity of demand, \( MC = C(Q, Z; \beta) \) - marginal costs.

**4) Structural Demand Models** (Dick 2002), according to which consumers choose for a particular bank based on prices and bank characteristics using consumer’s utility function:

\[
u_{ij} = \delta_{ij} + \epsilon_{ij} = p_{i}^{\alpha} \rho_{ij}^{\alpha} - p_{i}^{\alpha} \rho_{i}^{\alpha} + X_{ij}^{\beta} \xi_{ij}^{\beta} + \epsilon_{ij} (17),
\]

where consumers \( c \) and banks \( i \) populate markets \( j \), \( u_{ij} \) - consumer utility, \( p_{i}^{\alpha} \) - deposit rate paid by bank \( i \) in market \( j \), \( X_{ij}^{\beta} \) - vector capturing \( k \) observed product characteristics for the (singular) product offered by bank \( i \) in market \( j \), \( \xi_{ij}^{\beta} \) - unobserved bank product characteristics \( \alpha^{\beta}, \alpha^{\gamma}, \beta \) - taste parameters.

A consumer \( c \) chooses a bank \( i \) in market \( j \) if and only if \( u_{ij} \geq u_{ij}^{\ast} \), for \( r = 0 \) to \( I \), with 0 the outside good and \( I \) the number of banks in market \( j \).

**5) Other structural models:**

A). Sunk-Cost Models (Vives 2000, Dick 2004, 2006), according to which a bank’s market share is sufficiently responsive to investments in information technology (quality investments). The outcome of this “competition through endogenous sunk costs” is that the number of “dominant” banks in the market remains approximately the same and that only the number of “fringe” banks will increase in market size.

B). Structural Models of Entry (Bresnahan and Reiss 1991, 1994; Cohen and Mazzeo 2003), according to which the entry decisions of potential competitors and the continuation decisions of the incumbent firms only occur in case these decisions are actually profitable. The entry decision hinges on the level of fixed costs, the nature of post-entry competition, and the (future) entry or continuation decisions of other firms.

Thus we have studied different models which help us to measure contestability in banking markets and explain the key factors of such bank competitiveness.

**IV. Conclusion**

Because of prudential regulation in banking the standard paradigm of perfect competition is not considered appropriate for banking as the free entry and exit of banks from the market is not possible. Thus competition in banking has some specific features which distinguish it from the competition in other industries.
The measurement of output of banks services has special problems because of the intangible nature of banking. That is why there are two different approaches to measure bank output: the production approach (bank outputs are measured by the number of deposit and loan accounts or the number of transactions per account - a flow produced per unit of time) and the intermediation approach (bank outputs are measured by the value of loans and investments - a stock, showing the given amount of output at one point in time). Thereby because of the fewer problems the intermediation approach is used in most bank productivity studies which introduce banks as rational economic agents that act based on its specific productivity and efficiency functions. In banking it is very important to identify special problems which arise from the equivocation in the production function of banks and are the reason of efficiency measurement modelling based on cost and/or profit function which could explain the result of banks. There are two approaches of testing X-efficiency in banking (parametric and non-parametric), its formulas can help to understand the following models of competition in banking sector.

There are different models which help to measure contestability in banking markets and explain the key factors of such bank competitiveness. Some of them study a bank as a traditional industrial organization: Structure-Conduct Model, Relatively Efficiency Model and Economies of Scale and Scope Model. The Structure-Conduct-Performance Model sees the key factor in a change in the market structure which affects the way banks behave and perform. Such model supports the government intervention in order to limit monopoly power and decrease market concentration. The Relatively Efficiency Model sees the key factor in more efficient than others which is reflected in greater output. Such model doesn’t support the government intervention into banking but mergers and acquisitions of banks should be encouraged in case when it increase relative efficiency and should be prevented when it increase market concentration. The Economies of Scale and Scope Model see the key factor in optimal output mix both in terms of size and composition.

However because of the one-way causality between market structure and performance such models are being criticized. That is why new approaches have emerged which study a bank as a New Empirical Industrial Organization and do not try to infer the degree of competition from “indirect proxies” employing a variety of alternative methodologies.

The Contestable Banking Markets Model sees the key factor of bank competitiveness in a contestable market where players could gain an appropriate market share offering lower prices nevertheless the number of players on that. To test it such model offer the Panzar-Rosse Statistics – the measurement technique of market power through investigating the extent to which changes in factor input prices are reflected in equilibrium industry or bank-specific revenues. Such model doesn’t need government intervention because banks offer their products at marginal costs maximizing customers’ surplus. But it has some restrictions connected with constant improving information technologies which shorten the period between the entry of new player and his decision to exit of the market.

The Generalised Linear Pricing Model tests for the degree of competition in the banking market through a benchmark for a perfectly competitive rate (Libor Index - the rate banks quote each other for overnight deposits and loans) against which deposits and loan rates can be compared. The Conjectural-Variations Model sees the key factor of bank competitiveness in choosing its output takes into account the “reaction” of rival banks. The Structural Demand Models – in prices and bank characteristics which are reflected in consumer’s utility function. The Sunk-Cost Models - in investments in information technology (quality investments). The Structural Models of Entry – in profitable entry decisions of potential competitors and the continuation decisions of the incumbent firms only occur in case these decisions are actually profitable.

According to this we can define bank competitiveness as the possibility of bank to do its business efficient in order to gain profitable realization of its products in a competitive market.

V. REFERENCES

5. Heffeman, Sh. (2005) Modern Banking, John Wiley & Sons Ltd, Chichester, XIX.