4. SIZING UP PERFORMANCE MEASURES IN THE FINANCIAL SERVICES SECTOR

Jacob A. Bikker[1]

ABSTRACT

The adequate performance of banks, insurers and pension funds is of crucial importance to their private and business customers. The prices and quality of financial products sold by such entities are largely determined by operational efficiency and the degree of competition in the markets concerned. Since efficiency and competition cannot be observed directly, various indirect measures in the form of simple indicators or more complex models have been devised and used both in economic theory and in business practice. This paper demonstrates that measuring the performance of financial institutions is no simple matter and that indicators differ strongly in quality. It investigates which methods are to be preferred and how by combining certain indicators stronger measures may be developed. These measures are then subjected to a predictive validity test.

4.1. INTRODUCTION

This paper addresses the question how well financial institutions are performing in providing their services to consumers and businesses, and how much we know about that. Various performance aspects cannot be observed directly whereas they are economically important. While stockholders will view performance in terms of the profits made on their behalf, whether or not adjusted for risks taken, this paper focuses on performance in a broader sense, that is, the contribution financial institutions make to the common wealth, on behalf of consumers and businesses. They will be mainly interested in whether financial products are not too expensive and whether the quality is sufficient. This raises the issue of, on the one hand, the efficiency of financial institutions (*i.e.* whether unnecessary costs are made in bringing a product to market) and, on the other, the level of competition in the relevant markets (*i.e.* whether profit margins are not unnecessarily high). Since efficiency and competition cannot be observed directly, they have to be measured in an indirect way. If a cut in mortgage rates by one bank, for instance, is promptly copied by all its competitors, then this is a sign of competition – even if it does not enable us to distinguish between a little competition and strong competition. Yet the price and quality of other banking services such as

[1] The author is grateful to Laura Spierdijk for valuable comments and suggestions.

investment consultancy or payment services are much harder to determine, making competition far more difficult to measure. Difficulty in determining prices and quality levels, incidentally, is a widespread phenomenon in financial products markets. A recent example in the Netherlands is the investment-linked insurance policy, popularly known as 'robber policy' (*woekerpolis*). The fact that consumers find it hard to pick such a product on the basis of price and quality takes away the disciplinary influence of the customer and weakens competition. This problem inhabits many of the products of banks and insurers (Bikker and Spierdijk (2009a)).

There is another kind of performance that works in the interest of consumers, but does so in the long run. It is the reliability of a financial institution in terms of solvency and of whether customers can be sure to get their money back. Now that the subprime mortgage and liquidity crisis has engulfed us all, the amount of risk banks take in carrying on their business is a focal point of attention. Although this long-term performance is also affected by competition and efficiency, this paper concerns itself solely with the more palpable short-term performance exhibited in quality services and affordable prices.

Banks of course play a crucially important role in the economy because of their core products: loans to businesses and for house-purchase. Hence competition and efficiency in banking are also highly important: high quality at low cost boosts welfare. Competition is also important for adequate monetary transmission, which is the speed at which policy interest rates set by central banks pass through to bank interest rates (see Table 1).

Table 1: Importance of Competition in Banking

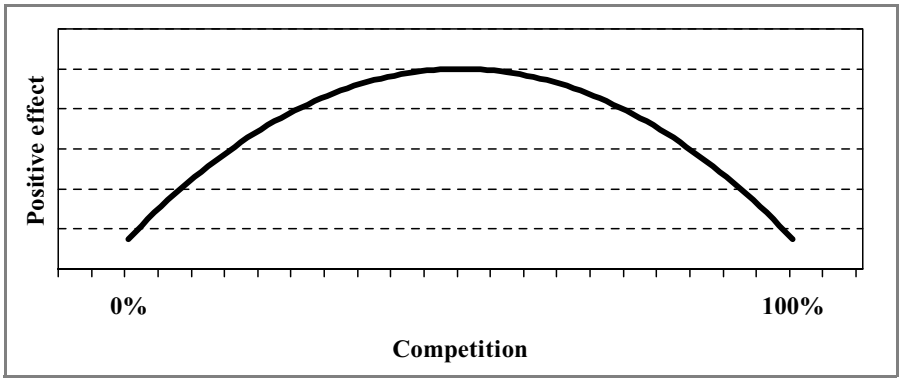
Welfare-enhancing for consumers and businesses
Reinforces monetary policy
Inverse U-shape relationship with:
– innovation
– solvency
– financial stability
– accessibility of the banking system to customers

Competition also affects financial innovations, banks' financial health, financial stability and the accessibility of banking services to customers – with accessibility meaning the extent to which small and medium-sized businesses have access to affordable financing. For all these four factors, the relation to competition is represented by a so-called inverse U-shape (see Figure 1). Promoting competition enhances these factors up to an optimum, whose position is uncertain. Stronger competition beyond the optimum has a counterproductive effect on these factors.

To give an example: when competition is very strong and excess profits dwindle, banks will find it hard to build extra buffers to protect them from adverse shocks. Healthy competition, in this sense, is better than fierce competition.

So what do banks, scientists and supervisors actually know about important variables such as competition and efficiency in the banking system? This paper will establish that, perhaps to our surprise or disappointment, we know far less than has often been taken for granted.

Figure 1: Positive Effect of Competition on Innovation, Financial Health and Accessibility of the Banking Industry and on Financial Stability



In practice, highly simplified approximations have been used to represent competition or efficiency, such as the concentration index or the cost-to-income ratio. While some indicators have been used without challenge in even the most highly-ranked scientific journals, they are in fact too primitive in nearly every case and not very reliable.

Better than such simplified proxies are theoretically founded models that attempt to estimate competition and efficiency for a particular country[2]. How well have these models been doing? This paper shows that the consensus between even the best-founded models is surprisingly weak. In other words, different methods lead to sometimes widely different results for the same country. This brings us to the central problem addressed by this paper: how far does the sounding rod of our measuring methods reach? And what can we do to reach just a little deeper?

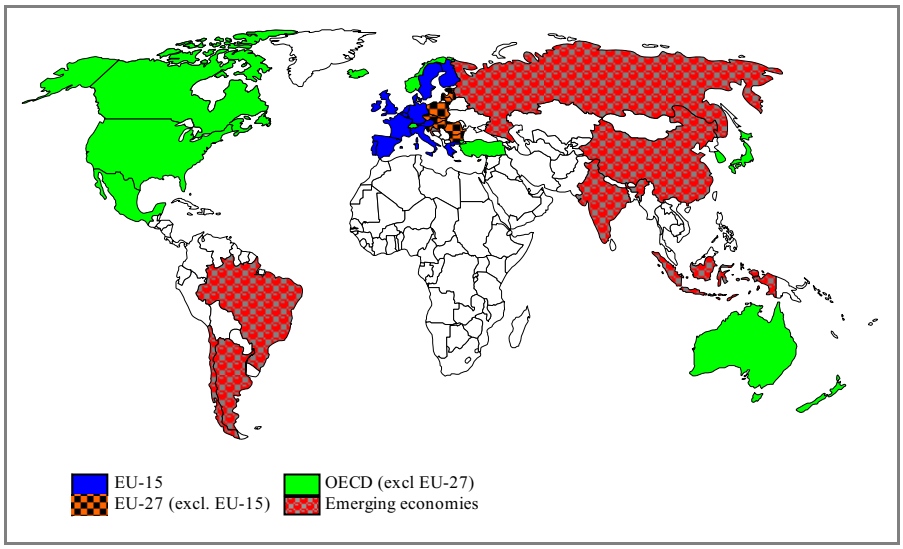
[2] Or for a particular bank. This paper considers country estimates.

4.2. PERFORMANCE MEASURES FOR FINANCIAL INSTITUTIONS

As a first step toward a closer analysis, about 20 methods were used to measure banking competition and efficiency for the most important 46 countries[3]. These countries comprise the old and new EU countries (in Figure 2 these are darkly shaded and chequered, respectively), the other OECD countries (light shading) and emerging markets (polka dotted). Together, they account for 90% of global GDP.

All 20 simple approximations and model estimates of competition will from now on be referred to as indicators. Five types of performance indicators are distinguished (see Table 2). Apart from competition and efficiency, these are costs, profit (margin) and market structure.

Figure 2: Countries Examined by Category



[3] For the list of these countries, see Bikker and Bos (2008), Table 9.1. Where competition is concerned, one country, Romania, was left out due to data issues.

Table 2: Indirect Performance Indicators for Financial Institutions

Performance indicators	Correlation with competition	Indicators represented as
Efficiency	Positive	Cost X-efficiency Profit X-efficiency Scale economies Scope economies
Costs	Negative	Cost-to-income ratio Cost margin Total costs/total income
Profit	Negative (?)	Return on capital Return on assets Net interest margin
Market structure		Number of banks
– number of banks	Positive	Per capita number of banks
– concentration	Ambivalent	HHI, C_3, C_5, C_{10}

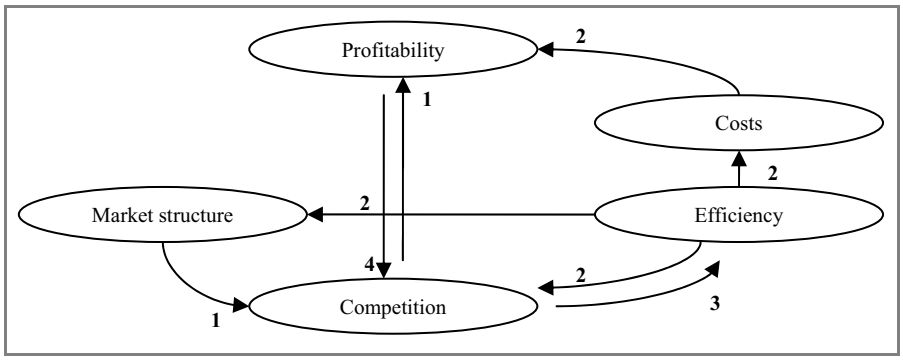
4.2.1. Mutual Relationships

Various theoretical relationships exist between the several types of performance. Figure 3 illustrates this with some examples. The classic structure-conduct-performance (SCP) theory holds that market structure determines competitive conduct and hence profits (referred to by the figure '1')[4]. For instance: high bank concentration leads to less competition and hence to higher profits. According to an alternative paradigm, the efficiency hypothesis, more efficient banks increase their market share by pushing less efficient competitors from the market (Demsets (1973)). More efficient banks will translate lower costs into either increased profits or price reductions – the latter in order to improve their competitiveness and increase their market share (indicated by a '2' in Figure 3). Efficiency thus is not an effect but a determinant of market structure[5]. It has been generally assumed that competitive pressure forces banks to become more efficient (indicated by a '3'). Hicks (1935) assumes as much, proposing, in his 'quite life' hypothesis, that monopoly will reduce the pressure towards efficiency. Finally, excess profits enable banks to lower their prices and become more competitive in order to increase their market share (indicated by a '4').

The strong intertwinement between variables in Figure 3 explains why market structure, costs and profitability are often used as proxies for competition and efficiency. At the same time, however, the figure underlines the fact that the measures concerned reflect quite different characteristics of banks and their markets.

[4] See Bos (2004) for an overview and a critical analysis.
[5] Depending on the ambition of efficient firms to expand their market share.

Figure 3: Relations Between Market Structure, Competition, Profitability and Efficiency



Explanation: Relations according to the SCP paradigm are indicated by the figure 1, those according to the efficiency hypothesis by the figure 2. Relations according to the 'quiet life' hypothesis (and its reversal) are marked by the figure 3, while the relation following from a general principle is indicated by 4.

4.2.2. Correlation with Competition

Before the indicators can be used, it must be established whether the correlation (across all countries) with competition is positive or negative[6]. Figure 3 shows that efficiency is positively correlated to competition (for stronger competition leads one to expect higher efficiency) and, for the same reason, that costs are negatively correlated with competition (in other words, stronger competition leads to cost cuts; see Table 2). Also, competition is likely to reduce profits. This argument is not entirely cogent, however, because competition may also affect profit in a positive sense via cost reduction. Hence the question mark in Table 2.

Where the notion of market structure is represented by the number of banks, a positive correlation with competition is usually assumed: the presence of more banks implies more opportunity for competition. Concentration, indicating mainly the dominant position of a small number of banks, may indicate low competition, because banks may use this to collaborate. A more dynamic interpretation is that such concentration may, on the contrary, be an indication of competition because consolidation may have been enforced by circumstances. Therefore concentration is an ambivalent indicator.

4.2.3. Models and Indicators Used

Initially, five models were used to estimate competition (see Table 3). The Lerner index uses profit margin as an indicator of market power (De Lange van Bergen

[6] Abstracting from causality. In some cases there are more theoretical connections, whereas different empirical results have been obtained. A final choice is made in all cases.

(2006)). The SCP model measures the influence of market structure on profits via an assumption of competitive conduct. Market structure, here, is approximated by the concentration index. The Cournot model is built along analogous lines, but instead of looking at the structure of the market as a whole, it regards the conjectural variation of individual banks[7]. Taking market share of the individual firm as a measure of market structure, the Cournot model aspires also to capture part of asymmetrical market structures, differences in cost structures and collusive behaviour. The Boone indicator measures how efficiency, through increased market shares, is rewarded by higher profits (Bikker and Van Leuvensteijn (2008); Boone (2004, 2008); Van Leuvensteijn *et al.* (2007, 2008)). The Panzar-Rosse model measures to what extent input and output prices move in step (as they would under perfect competition) or out of step (indicating monopoly or a perfect cartel)[8]. Other models in the literature (*e.g.* Bresnahan, Iwata) require data sets that for most countries are simply lacking, while estimations also present high practical barriers (Bikker (2003)). Table 3 shows how the different models simulate different aspects of competition.

Table 3: Competition Models

Model	Underlying concept
Lerner index	Profit margin indicates market power
SCP model[a]	Effect of market structure (concentration) on profit through competitive behaviour
Cournot model[b]	Effect of market structure (market share) on profit through competitive behaviour
Boone indicator	Degree in which efficiency is rewarded in the form of higher profits through increased market shares[c]
Panzar-Rosse model	Correlation of input prices and income (revenue)

a. Based on, respectively, the market shares of the largest three banks (C_3) and the Herfindahl-Hirschman concentration index (HHI) as measures of market structure.
b. Based on the market share of the individual bank as a measure of market structure, as an indicator of asymmetrical market structures, differences in cost structures and collusive behaviour.
c. Based on the efficiency hypothesis.

For the efficiency indicators, cost and profit X-efficiency as well as scale and scope economies were estimated through a model (see Table 2). Costs are represented by the cost-to-income ratio and the cost margin, while profit is proxied by return on capital or return on assets (RoA) and by net interest margin (NIM). In the case of market structure, the number of banks, the per capita number of banks and a number of concentration indices are also incorporated[9].

7 Conjectural variation is the degree to which a bank in setting its prices and total production quantity in a business area is aware of its dependency on other banks' behaviour in that area.
8 See Panzar and Rosse (1987).
9 For the exact definitions, see Table 16.1 in Bikker and Bos (2008). Concentration indicators are discussed in Bikker and Haaf (2002a).

In all cases this analysis was based on the banking market as a whole, without regard to product differences. It has been argued against this that the situation as regards competition, for instance, may vary depending on the market segment. Competition in the mortgage lending market is likely to be much stronger than in the investment counselling market. This is justified criticism: competition may vary from product to product or even from one location to another. However, for most products there are insufficient data available to perform analyses at the product or location level, with a few exceptions[10]. Where approximations for competition and efficiency are used in the economic literature, this is almost invariably done for banks as a whole, so on the highest level of aggregation.

Since all models were estimated on the basis of a single dataset, different outcomes may not be attributed to data differences. The dataset covers a ten-year period (1996-2005) and was obtained from Fitch IBCA's BankScope and from the OECD[11].

4.3. CRITICAL APPRAISAL OF THE INDICATORS

This section appraises the indicators presented above against three different criteria: first, two statistical norms – mutual correlations and the principal component analysis[12] and an economic interpretation. Finally, the variation across countries is explained from economic theory.

4.3.1. Correlations

How do the different indicators found correlate to each other? Table A.1 in the appendix shows the correlation coefficients between 14 currently used indicators for 46 countries[13]. A correlation between two variables indicates parallel move-

[10] Bikker and Haaf (2002b) and Bikker *et al.* (2006b) use the Panzar-Rosse model to disaggregate by bank size, thus going some way towards a breakdown by market type (international vs. local), client type (large corporation vs. medium and small-sized businesses) and product type (wholesale vs. retail). Van Leuvensteijn *et al.* (2007) estimate competition in just the lending market.

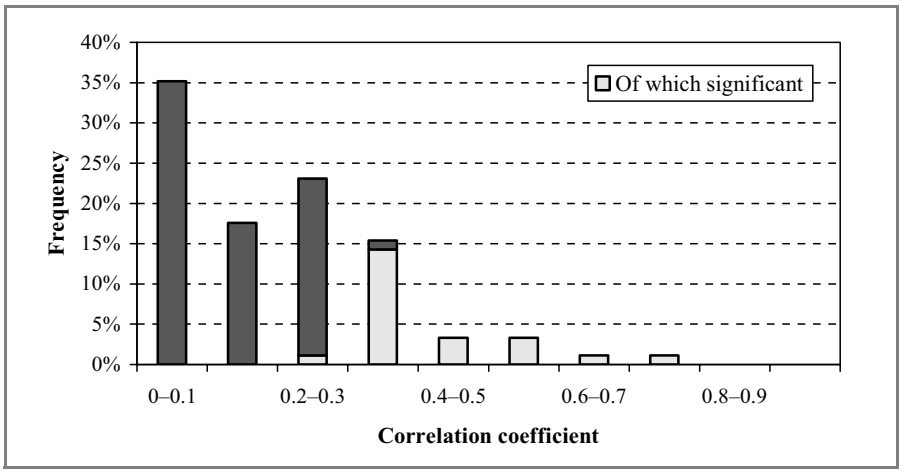
[11] De data on individual banks' balance sheets and profit and loss accounts that were used by the five competition measuring models and the models to measure X-efficiency were obtained from BankScope. The dataset contains data on 13,000 private and public banks publishing more or less standardised annual accounts which permit comparison between the different accounting systems. The data underlying the profit and cost indicators for the OECD countries were obtained from the OECD (2000, 2002, 2004). Those data coincide with those used by Bikker and Bos (2008) and are discussed more fully there. The data on concentration indices for all countries and those underlying the profit and cost indicators for the sixteen non-OECD countries were calculated on the basis of the banks from those countries that figure in BankScope. Selection rules were applied to the latter set in order to eliminate banks in unusual circumstances (*e.g.* holdings and banks undergoing a start-up or winding down process). See Bikker *et al.* (2006a).

[12] A third statistical method might have been regression analysis. However, the use of this is doubtful given the strongly endogenous nature of (almost) all variables used. A counterexample is Koetter *et al.* (2007).

[13] All analyses for 46 countries were made without the Lerner index. Lerner index analyses were performed for 23 countries, but are not discussed here since the index turns out to be significantly correlated only with the Boone indicator. Table A.1 is part of a larger correlation matrix, because the total number of variables investigated was larger than 14.

ment, without regard to any original (causal) connection. Figure 4 summarises these findings as the frequency distribution of the correlations found.

Figure 4: Frequency Distribution of Correlation Coefficients Between Indicators



Explanation: The graph presents the 91 correlations between the 14 indicators used: the Boone, Panzar-Rosse, SCP and Cournot models, cost and profit X-efficiency, return on assets or on equity, cost-to-income ratio, total cost to total income ratio, net interest margin, cost margin, the number of banks and the top 5 banks by market share, C5 (see Table A.1). Lighter shading refers to the 22 correlations that are significant at the 5% significance level.

Evidently, most correlation coefficients are below 0.5: apparently, indicators tend to be only moderately correlated to each other. This underlines the fact that each single indicator provides at best a rough indication of competition, which is certainly not very accurate at the country level. The lighter shading indicates correlations that are significant at the 95% confidence level – the upper fourth part of all results. The number of significant correlations, at one in four, is not very high. However, they all have the right – meaning: theoretically expected – sign, except for five correlations involving ambivalent indicators whose sign depends on which of the several theoretically possible relationships is dominant. The fact that all other 17 significant correlations bear the right signs without exception is an indication that the indicators behave (roughly) in accordance with the theoretical framework and hence are not too much distorted by *e.g.* definition or measurement issues.

4.3.2. Principal Components Analysis

Another statistical technique is principal components analysis or PCA[14]. To the extent possible, this method attempts to represent the variation across the countries within a set of correlated variables using a few variables called principal components. PCA makes it possible to investigate to what extent the indicators reviewed might all be explained by just a few factors or, in other words, to what extent they overlap. The more successful the analysis, the more similar to each other the indicators would be. Even more important is the possibility to interpret the principal components (PCs) and to see whether they might represent recognisable elements of our performance measures. It would be nice, for instance, if one of the PCs represented competition, another one efficiency and the third one profitability. This way, each PC could, so to speak, filter information from the indicators and represent it in compact form.

Table 4 shows the outcome of an analysis (after the so-called varimax rotation for ease of interpretation) with twelve indicators, selected so as to minimise overlap between the indicators considered. Also, the indicators are spread as equally as possible across the categories competition, efficiency, profitability, et cetera[15]. The shading indicates for each column (*i.e.* for each principal component) the highest factor or component loading(s). Thus we may infer that the first principal component represents mainly cost and profit margins and profit inefficiency[16]. The second one has the highest factor loading at cost efficiency, while the third one has its highest factor loading at (three out of four) model-based competition measures, and again at the HHI concentration index. Apparently, this third factor comprises information on competition. Moreover, the signs of each factor loadings are correct – that is to say, in accordance with theoretical expectations[17] – so that this PC ought to present a reliable summary of the information content of these competition indicators.

[14] PCA is a multivariate statistical technique that defines, for a large number of observed variables, a smaller number of underlying series. As a statistical method, PCA is nearly identical to factor analysis. Apart from data reduction, PCA aims to provide an understanding of the dataset's structure.

[15] If the selected indicators are varied a bit, the outcome of the PCA will change as well. Typically, the first PCs may usually be interpreted as profit, efficiency and competition – though not always in that order. In some cases, costs appear in combination with profits, while in others they are coupled with efficiency.

[16] Note that competition depresses both costs and profits.

[17] As competition grows, the *H*-values of the Panzar-Rosse model will also rise, whereas the Boone indicator and the coefficients in the SCP model and the Cournot model decline.

Table 4: Factor Loadings for the First Five Principal Components (PCs)

	Factor loadings[a]					Explanation[b]
	PC1	*PC2*	*PC3*	*PC4*	*PC5*	
Panzar-Rosse model	-0.20	0.18	0.80			0.72
Boone indicator	0.20	0.30	-0.79			0.76
SCP model	-0.80	0.18				0.67
Cournot model	0.18	-0.23	-0.63	-0.42		0.66
Cost efficiency	-0.13	0.81	0.13	0.11		0.70
Profit efficiency	0.84	-0.24				0.76
Return on Assets	0.79	0.16	-0.27	-0.24		0.79
Cost to income ratio	0.26	-0.60	0.60	0.14		0.81
Net interest margin	0.84	-0.18	0.18			0.77
Number of banks	-0.20	0.13	0.12	0.85	-0.12	0.81
Cost margin	0.85	-0.23	0.12	-0.13		0.81
HHI	0.19	0.13	-0.85	-0.14		0.79
	Explanation of variance per PC					*Total*
	0.19	0.17	0.13	0.15	0.12	0.76

a. A factor loading may be regarded as the coordinate of an indicator on a PC in a coordinate system. In the case of orthogonal components (i.e. forming a right angle), the factor loading of a variable vis-à-vis a component equals the correlation between that variable and that component.
b. Explanation of the variance of the indicators based on the first five PCs (equals the sum of squared factor loadings for each variable across the five PCs).

Explanation: The shading indicates the highest factor loading for each column (that is, PC).

The last line of Table 4 shows that the first PC explains almost 20% of the variance in the indicators, falling gradually to 12% for the fifth PC, so that the first five PCs together explain 76% of the variance. Thus less than half the PCs explain three-fourths of the variance in the indicators. Apparently, the indicators do contain common elements (especially 'competition'), but also many specific ones (profit, efficiency, concentration and further refinements such as RoA and NRM).

4.3.3. Economic Interpretation

What, now, is the economic significance of the indicators, or what are their countryspecific values? The answers to these questions are found, for the present estimates of country-level competition and efficiency, in comparing the results to available other sources of a more intuitive or anecdotal nature, or that relate to specific subsegments or to competition in other sectors. However, there is not much contrastive material around. In practice, there seems to be a degree of consensus to the effect that Anglo-Saxon countries such as the USA, the UK and Ireland are highly competitive. Another *expert view* is that competition in Southern Europe, by contrast, is very modest as a result of lagging development, exem-

plified by insufficient consolidation and low cost-sensitiveness in bank clients. France and Germany are also (with Italy) supposed to be less competitive owing to strong public interference and inadequate consolidation. Very recently, we have seen strong government interference with banks in many countries, in response to the financial crisis – good for solvency but bad for competitive conditions and therefore, one hopes, temporary. For Germany, stricter adherence to supervisory rules, financial conservatism and an extensive branch network are mentioned. Another universally accepted truth is that competence is stronger in developed countries than in emerging economies, with the least developed countries bringing up the rear. Table 5 presents the country ranking according to the 'expert view'.

Table 5: Competitiveness Ranking of EU Countries: Expert View vs. Empiricism

	Expert View		Empiricism (indicators)
1	UK/USA/Ireland	1	Germany/France
2	Western Europe	2	UK/USA/Ireland
3	Germany/France	3	Other EU-15 countries
4	Southern Europe	4	Central & Eastern Europe
5	Central & Eastern Europe		

Various indicators produce diverging results for the same countries, because they reflect different aspects of competition and also because estimation errors or faulty data distort the result. But there is something else, which is that the outcome suggests the above generally accepted country ranking is, in fact – or at least according to our estimates – simply wrong. Germany, which is deemed by many to be low on competition, gets good marks for all our criteria: low cost, low profit, high competition, high efficiency – and as measured by nearly all indicators. And a very similar story applies to France. Some Southern European countries live up to their underdeveloped image, yet according to many indicators, Italy – and to some extent Spain – do not. Conversely, the performance measures for the USA, the UK and Ireland are less than convincing. Although competition estimates for these countries are favourable, their cost levels (and cost inefficiencies), interest margins and profits are exceptionally high, which is hard to reconcile with a competitive climate. Table 5 shows that according to the indicators as measured across 1996-2005, Germany and France take the lead over the Anglo-Saxon countries.

Whereas the original purpose of the above comparison was to use the 'generally accepted truth' as a benchmark for the indicators, the outcome suggests the reverse, *i.e.* the urgent need to adjust the expert view.

4.3.4. Causes of Country-level Deviation Among Indicators

What is it that causes various measures to reflect somewhat different phenomena for each country? There are three main explanations. First, we are dealing with different concepts: although mutually correlated, the indicators do in fact measure different things: competition is not the same thing as efficiency, which in turn differs from profitability et cetera. Secondly, there are definition issues: each definition of (for instance) efficiency reflects a different aspect of the concept. And finally, imperfections in the data also play a role.

Definition issues also figure in the models that measure competition. Using a standard model of a profit maximising bank under a regime of oligopolistic competition, one may derive that the theoretical model of competence is as follows (Bikker and Bos (2005, 2008)).

$$\textit{Profit margin} = (\bullet 1/\mu)\ \textit{HHI}\ (1 + \lambda) \tag{1}$$

Profit margin is assumed to reflect competitiveness: the more market power, or the less competition, the higher profits will be. The μ parameter indicates the price elasticity of demand: the more sensitive consumers are to changes in the prices of bank products, the stronger competition will be. *HHI*, the Herfindahl-Hirschman index of concentration, describes market structure: more banks make for more competition, while a market with few large banks weakens competition. The conjectural (or assumed) variation, λ, indicates how banks will respond to production volumes and prices of other banks. This parameter becomes higher as competition gets stronger. Equation (1) may also be derived at the firm level where, applied to bank *i*, it reads:

$$\textit{Profit margin}_i = (\bullet 1/\mu)\ MS_i\ (1 + \lambda_i) \tag{2}$$

where MS represents market share. Bikker and Bos (2005, 2008) have demonstrated that existing competition models may be derived from these two, except that they invariably incorporate only one or two of the three components, thereby neglecting one or two others. The SCP model, for instance, assumes that μ and λ in equation (1) are constant (or that $(1 + \lambda)$ may be approximated by HHI). The same goes for Cournot, albeit at the bank instead of the country level (see equation (2)). The Boone indicator is estimated as the μ in equation (2) and assumes λ_i constant. These differences in *a priori* assumptions contribute to the variation in competition estimates. The Lerner index and the Panzar-Rosse model base themselves on the (full) profit margin at the firm level. In the case of the Lerner index, there is the problem that marginal costs have to be estimated, while with Panzar-Rosse the translation from theoretical to empirical model may have a disturbing effect.

4.4. WHAT INDICATORS CAN DO

In the preceding paragraphs competition it has been shown that competition indicators should not be applied indiscriminately. Time to investigate what information value the indicators do have and whether there is, in fact, a reliable way to gauge competition. In order to find this out, we will be concentrating on three aspects: economic interpretation (again), predictive validity and a bundling of all information into a single index.

4.4.1. Economic Interpretation

To see whether any clear structure lies buried inside the data, Table 6 presents the estimates of the average cost and profit X-efficiency, costs (averaged across the three cost indicators) and profitability (averaged across the three profit indicators). The table juxtaposes three types of countries (*viz.* (1) Western Europe and other highly industrialised countries, (ii) emerging economies and other OECD countries, and (iii) Eastern and Central Europe) with efficiency, broken down into five classes in descending order from high to low efficiency countries. Every cell in the table contains the number of countries in that bracket. The table shows a diagonal pattern (see shading). Apparently, the efficiency of banks in the highly developed industrial countries is clearly better than that of banks in emerging countries, while banks in the post-transitional economies of Eastern and Central Europe come out as least efficient. It follows that there is a correlation between efficiency and degree of economic development.

A similar pattern from high to low is to be found, for the same reason, when countries are classified by cost levels or profitability, but that the other way around (from low to high) as high efficiency corresponds to low costs and low profits (see the shaded diagonal in, respectively, Table 6.B and 6.C). In the developed countries, where costs are lower, profits are also lower, whereas costs and profits are higher in the transition countries. It is tempting to ascribe this phenomenon to stronger competitive pressure. However, a similar classification does not show an unequivocal pattern for competition. Other investigations have shown that competition in industrial countries is, by contrast, slightly weaker, probably owing to a higher proportion of products such as investment counselling and services and options, where competition is far less energetic than on deposit taking and lending (Bikker *et al.* (2007). In time, the share of advisory and other services continues to increase, further weakening competition (Bikker and Spierdijk (2008)).

Table 6: Distribution of X-efficiency, Costs and Profitability Across Countries

		Western Europe and other industrialised	Emerging economies and other OECD	Eastern and Central Europe
		A. X-efficiency		
High	9	8	1	
	9	7		2
Medium	9	3	3	3
	9	5	3	1
Low	9	2	2	5
	45[a]	25	9	11[a]
		B. Costs		
Low	9	9		
	8	6	1	1
Medium	10	6	2	2
	9	4	3	2
High	10		3	7
46		25	9	12
		C. Profitability		
Low	9	7	1	1
	9	6	2	1
Medium	9	6	2	1
	10	6	1	3
High	9		3	6
	46	25	9	12

a. The X-efficiency of Romania could not be estimated due to insufficient data.

4.4.2. Average Ranking

In situations where measuring is problematic, a good solution may well be to take the average of several estimations. This is a well-known and often-used strategy in forecasting: the combination of several forecasts does better than each forecast separately. This strategy was also applied to the set of estimates and indicators discussed above. A per-country average of several competition level indicators was used. Because the units of expression of these indicators cannot be compared, instead of values, ranking orders were averaged[18]. For this exercise, eleven measures were selected in such a way that there was as little overlap between them as possible. Wherever the overlap between two measures was substantial, one vari-

[18] The third principle component 'competition' as presented in Table 4 is an alternative index, which may be viewed as a weighted (by factor loadings) average of the original normalised series.

able was left out[19]. The eleven eventually selected measures are: Boone indicator, Panzar-Rosse model, SCP model, Cournot model, cost X-efficiency, return on assets, cost-to-income ratio (C/I), total cost to total income ratio, net interest margin (NIM), cost margin (CM) and market share of the top 5 banks (C_5).

Table 7: Correlations Between the Indicators and the Index

Indicators	Correlations	Significance	Status	Index component
Boone indicator	-0.14			Yes
Panzar-Rosse model	0.33	**		Yes
SCP model	-0.05			Yes
Cournot model	-0.42	***		Yes
Profit efficiency	0.37	**	Amb.	
Cost efficiency	0.53	***		Yes
Return on capital	-0.30	**		
Return on assets (RoA)	-0.50	***		Yes
Cost-to-income ratio (C/I)	-0.42	***	Amb.	Yes
Total cost to total income ratio	-0.20		Amb.	Yes
Net interest margin (NIM)	-0.63	***		Yes
Cost margin (CM)	-0.58	***		Yes
Number of banks	0.51	***		
Concentration index C_5	-0.37	**		Yes

Note: Two (three) asterisks indicate a confidence level of 95% (99%). Shading indicates expected positive correlation. (Only where there is ambivalence is there no a priori expectation. Amb. stands for ambiguous.

Table 7 (*i.e.* the last column of Table A.1) presents the correlations between the 'average ranking', referred to from here on as 'Index', and the underlying variables. Remarkably, 11 of the 14 measures are significantly correlated with the Index, of which 7 at the highest confidence level of 99%[20]. Figure 5 shows, moreover, that correlations with the Index are far stronger than those between pairs of indicators.

Reassuringly, all 14 correlations have the correct (theoretically expected) signs[21], which is, of course, especially significant in the case of the nine significant and non-ambivalent variables: Panzar-Rosse model, Cournot model, cost X-efficiency, return on assets/capital, NIM, CM, number of banks and C_5. Apparently

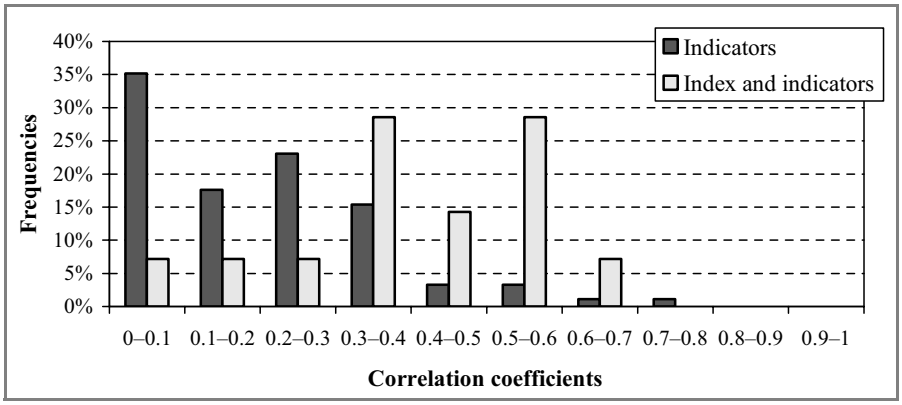
[19] Cost-based or profit-based scale economies were also disregarded because they show little variation across the countries and because of its ambivalent relation to competition.

[20] For the indicators included in the Index, a modicum of correlation with the Index is to be expected, of course. While for some indicators (Boone indicator and SCP model) this does not lead to significance, other indicators show significant correlation without being included in the Index (*e.g.* profit efficiency and number of banks).

[21] The correct sign is negative (owing to the selection made in constructing the Index, because most indicators correlate negatively with competition, see Table A.1), except in certain cases (shading).

there is, after all, an overall concept of 'competition', which is present in nearly every indicator and is reflected reliably and unequivocally in the resulting Index[22].

Figure 5: Frequency Distribution of Correlations Between Indicators and Index



Note: Dark shading: frequency distribution of 91 correlations between indicators; light shading: frequency distribution of 14 correlations between indicator and Index.

Now that an adequate measure of competition has been found in the Index, it is possible to tell which of the simple indicators, all things considered, does best. Table A.1 shows that the net interest margin and its relation, return on assets, are the most successful (overall) performance measures[23]. When the focus is entirely on competition, Panzar-Rosse or Cournot are more satisfactory.

Finally, it should be noted that this ranking-based Index is strongly and significantly (and in declining degrees) correlated with the first three principal components of Table 4, which are weighted averages of he original indicators. Both the Index and the principal components aim to present as much of the indicators' information content as possible in summary form.

[22] A corollary result is that the ambivalent variables are now signed, so as to make clear which relation prevails in practice. In the case of profit efficiency the influence of cost efficiency dominates that of the use of market power. The cost-to-income ratio and the total cost to total income ratio turn out to do well as indicators of efficiency, with the enumerator (costs) dominating the denominator (income) in determining the ratio.

[23] In earlier analyses across a smaller number of countries, using a differently composed set of indicators (Bikker and Bos, 2008) or covering other periods (Bikker and Bos, 2005), the net interest margin and the return on assets also came out on top.

4.4.3. Predictive Validity Test

There is another way to test the measures considered, which derives from the psychometric, sociological and marketing literature: the so-called predictive validity test[24]. The predictive validity test is based on the idea that a constructed variable – such as a survey question – must be correlated to the (subsequently) observed variable if it is to be a useful predictor. With some adjustment the indicators in the present analysis could be subjected to the following 'informative validity test'. The test is based on a model in which competition depends on economic variables or, conversely, where an economic variable depends on, among other things, competition. In such a model each of our indicators might be used as a proxy for competition to see whether it is both significant and (according to theory) correctly signed. If it is, one may conclude that the indicator's relevant information content prevails without its pattern being disfigured by the inherent noise.

Such tests occur frequently in the literature, if implicitly, because indicators are usually employed without much ado as competition measures. Examples of this are the SCP and the efficiency hypothesis literatures where concentration and market share, respectively, have been blithely cast in the role of competition. But there are many other fields of study where competition comes into play[25]. As an *ex-post* test the literature is not a reliable source, since less welcome test results are more likely to be disregarded by authors or else to be rejected by journals.

Below are three examples of such informative validity tests. A model-based measure of competition is the H-value from the Panzar-Rosse model which has been estimated for 80 countries. Next, it is explained by means of a large number of carefully selected possible determinants of competition (Bikker *et al.* (2007)). The four (out of nine) determinants that are significant (even at the 99% confidence level), all turn out to carry the right sign (see Table 8). Apparently, the H statistic contains a great deal of – competition-related – information, so that it passes the present test successfully.

[24] Predictive validity is the term used if a test is observed before it can be compared to the realisation; 'concurrent validity' is applied in cases where observation is simultaneous. The latter term would be applicable if one indicator were to be validated against the other. This option is less useful in the present analysis owing to the endogenous nature of the indicators considered here.

[25] Some examples of this are given further below.

Table 8: Explanation of Bank Competition in 76 Countries (2004)

Variables	Coefficients	t-value	Significance
Concentration index C5	-0.001	-0.8	
Activity restrictions	-0.000	-0.7	
Log (Market cap./GDP)	-0.016	-0.4	
Log (per capita GDP)	0.011	0.3	
Real GDP growth	-0.023	-2.8	Sign.
Foreign investment index	-0.132	-3.2	Sign.
Regulation index	0.128	2.5	Sign.
EU-15	-0.129	-1.4	
Former planned economies	-0.435	-5.6	Sign.
R2, adjusted	0.82		

Source: Bikker et al. (2007).

Our second example concerns monetary transmission. It is assumed that as competition increases, bank interest rates will be lower and more closely aligned with market rates and the policy rates of the European Central Bank (ECB), so that competition reinforces monetary policy. Models for four types of lending in eight EMU countries[26] explain the spread between the observed four bank rates and the corresponding policy and market rates using competition in the lending market (Van Leuvensteijn *et al.* (2007, 2008))[27]. Competition was in this case measured by the Boone indicator, because it permits estimating competition in a partial market (*i.e.* the lending market). The competition measure carries the correct sign significantly for three out of the four lending rates (see Table 9). In the fourth case, the coefficient concerned is not significant. Also, a so-called 'Error Correction Model' shows that the response of all four lending rates to the market and policy rates is stronger, and hence more closely parallel, as competition increases. Again, the Boone indicator, with seven hits out of eight, seems to have passed the test[28].

[26] Austria, Belgium, France, Germany, Italy, the Netherlands, Portugal and Spain (1992-2004).

[27] An alternative model, the Error Correction Model, was unable to confirm decreasing spreads amid stronger competition. Apparently, this more complicated model is less capable of measuring the targeted alignment effect.

[28] In addition, the spread between two deposit rates and the corresponding market and policy rates is explained by competition on the lending market. It turns out that deposit rates tend to be lower the more competition there is on the lending market. Apparently, competition on lending is not a good indicator for competition in the deposits market. On the contrary: banks compensate for their loss of income as a result of competition on lending by offering lower deposit rates.

Table 9: Effect of Competition on Spreads between Bank and Market Lending Rates

	Effect of competition on spread (t-values)	Effect of competition times market rate on bank rates (t-values)
Mortgage loans	**-2.12	***4.29
Consumer credit	***-3.03	***3.21
Short-term corporate loans	***-6.72	***3.47
Long-term corporate loans	0.15	***4.48

Note: Two (three) asterisks indicate a confidence level of 95% (99%).
Source: Van Leuvensteijn et al. (2008).

A third example is that of a model which determines the influence of competition on a banks capital buffer (Bikker and Spierdijk (2009b)). On the one hand it seems self-evident that less competition should lead to higher bank profits, so that banks may add more money to their buffer capital. There is a clear trade-off here between the short-term interest of bank customers, characterised by high competition and low prices, and the long-term interest of financial safety, in other words, the certainty that you will get your money back. An alternative theory assumes, however, that when fierce competition erodes profit margins, banks will be inclined to take more risks and hold a smaller buffer. Also, amid strong competition, banks will be less inclined to invest in inquiries regarding their clients in order to reduce information asymmetry (Marcus (1984)). This, too, increases the risk for banks. In order to determine which effect is stronger, a model was estimated – on the analogy of work by Schaeck *et al.* (2006) and Schaeck and Cihak (2007) – where the capital buffer depends on variables including competition. Competition was once more measured using the Panzar-Rosse model, so that data are available for over 100 countries.

Estimations demonstrate that competition erodes banks' capital buffers, so that apparently, the theory claiming that 'weak competition leads to high profits and hence to large buffers' wins out in actual practice. The same holds if instead of the Panzar-Rosse competition measure the third principal component derived above (which according to the factor loadings indicated competition) has been applied[29]. Again it appears that measuring competition in practice yields plausible results.

[29] In fact, the Index turns out not to be significant if replacing the Panzar-Rosse measure.

4.5. WHAT DO THE VALIDATED MEASURES ACTUALLY MEASURE?

So far, this paper has been investigating how bank performance indicators do themselves perform as measures. Next, the question arises as to the banking industry's competition and inefficiency themselves. Earlier studies have tried to capture those variables. For the sake of comparison, two other financial sectors are also considered: insurers and pension funds. Little research has been done in the present area for these types of financial institution, while banking competition measurement has been underexposed in the literature.

This paper considers only estimates by methods whose results cover the same 0%-100% range, which permits the outcomes to be compared. Disregarding for now the many (almost insurmountable) problems besetting the business of measurements and comparisons[30], Table 10 presents several outcomes for scale economies, cost X-inefficiency and competition.

Unused scale economies cannot be present under strong or perfect competition. Estimated unused scale economies increase from banks (5%) via nonlife and life insurers (10% and 20%, respectively) to 36% for pension funds. Especially insurers and small pension funds could realise hefty cost savings through (further) consolidation. These outcomes reflect the degree of (overdue) consolidation per sector, and therefore in a sense a lack of competition. For under fierce competition, large-scale cost-saving opportunities would not go unexploited[31]. As has been observed many times, the inefficiency of banks and insurers is greater than their scale inefficiency. Bank competition, at 50% (world-wide), hovers halfway between monopoly and perfect competition[32]. In recent years bank competition has weakened somewhat (Bikker and Spierdijk (2008)). Among nonlife insurers, competition is considerably weaker, at 22%, than among banks (Bikker *et al.* (2008)). The conclusion is that there is a good deal of room for improvement in competition and efficiency within banks and, especially, insurance companies.

[30] The measurement of scale economies, for instance, is based on the variable 'output', which presents its own measurement issues for each sector.

[31] It should be noted that these scale effects also concern production structures. In all sectors, fixed costs are high and rising over time, while they are particularly high in pension funds, compared to variable costs.

[32] The competence measure H of the Panzar-Rosse model, measured across 100 countries, averages 0.50, exactly halfway between monopoly ($H = 0$) and perfect competition ($H = 1$).

Table 10: Competition Among Banks, Insurers and Pension Funds (per cent)

	Banks	Insurers		Pension funds
		Nonlife	*Life*	
Scale effects[a] (Int.)	5	–	–	–
Scale effects[b] (Nld.)	–	10	20	36
Inefficiency[c] (Int.)	18	–	–	–
Inefficiency[d] (Nld.)	18	–	28	–
Competition[e] (Int.)	50	22	–	–

a. Scale effects are defined as the average percentual savings on the operating costs of any additional production realised as a result of upscaling. The greater the unused scale economies, the weaker competition will be. Source: calculations by the author and Marco Hoeberichts.
b. Sources: Bikker and Van Leuvensteijn (2008).
c. Cost X-inefficiency. Source: Bikker and Bos (2008).
d. Sources: Bikker and Bos (2008).
e. Sources: H-values by Bikker et al. (2006a) and Bikker et al. (2008).

4.6. SUMMARY

While many indicators of competition between banks commonly used in economic literature and in practice do in fact measure something, they do not contribute much to our knowledge on bank performance. At the same time it has been established that with the help of appropriate indicators – or, even better, a *combination* of appropriate indicators – we could make a good deal of headway towards a better understanding of competition. The appropriate indicators contain sufficient information on competition to be able to function reliably as explanatory variables in a model where competition plays a dominant role. Finally, the analysis also revealed that some existing expert opinions on the relative competitiveness of (especially European) countries need to be thoroughly reviewed. Application of several indicators to banks, life & nonlife insurers and pension funds has consistently shown that there is a good deal of room for improvement on competition and efficiency in banks and, especially, insurers.

4.7. REFERENCES

BIKKER, J.A. and HAAF, K. (2002a), "Measures of competition and concentration in the banking industry: a review of the literature", *Economic & Financial Modelling* 9, 53-98.

BIKKER, J.A. and HAAF, K. (2002b), "Competition, concentration and their relationship: an empirical analysis of the banking industry", *Journal of Banking & Finance* 26, 2191-2214.

BIKKER, J.A. (2003), "Testing for imperfect competition on the EU deposit and loan markets with Bresnahan's market power model", *Kredit und Kapital* 36, 167-212.

BIKKER, J.A. and BOS, J.W.B. (2005), "Trends in competition and profitability in the banking industry: a basic framework", *Suerf Series* 2005/2.

BIKKER, J.A., SPIERDIJK, L. and FINNIE, P. (2006a), *Misspecification in the Panzar-Rosse model: assessing competition in the banking industry*, DNB Working Paper No. 114[33].

BIKKER, J.A., SPIERDIJK, L. and FINNIE, P. (2006b), *The impact of bank size on market power*, DNB Working Paper No. 120.

BIKKER, J.A., L. SPIERDIJK and FINNIE, P. (2007), *Market structure, contestability and institutional environment: the determinants of banking competition*, DNB Working Paper No. 156/Tjalling C. Koopmans Research Institute Discussion Paper Series No. 07-29, Universiteit Utrecht[34].

BIKKER, J.A. and BOS, J.W.B. (2008), *Bank Performance: A theoretical and empirical framework for the analysis of profitability, competition and efficiency*, Routledge International Studies in Money and Banking, Routledge, Londen & New York, 176 p.

BIKKER, J.A. and GORTER, J. (2008), *Performance of the Dutch non-life insurance industry: competition, efficiency and focus*, DNB Working Paper 164/Tjalling C. Koopmans Research Institute Discussion Paper Series No. 08-01, Universiteit Utrecht.

BIKKER, J.A. and SPIERDIJK, L. (2008), *How banking competition changed over time*, DNB Working Paper No. 167/Tjalling C. Koopmans Research Institute Discussion Paper Series No. 08-04, Universiteit Utrecht.

BIKKER, J.A. and VAN LEUVENSTEIJN, M. (2008), "Competition and efficiency in the Dutch life insurance industry", *Applied Economics* 40, 2063-2084.

BIKKER, J.A., MIRO, A. and SPIERDIJK, L. (2009b), *Competition in the non-life insurance industry: a global analysis*, mimeo.

BIKKER, J.A. and DE DREU, J. (2009), "Operating costs of pension funds: the impact of scale, governance and plan design", *Journal of Pension Economics and Finance* 8, 63-89.

BIKKER, J.A. and SPIERDIJK, L. (2009a), "Measuring and explaining competition in the financial sector" in *Proceedings on the G-20 meeting on Competition in the Financial Sector* (Bali), Bank Indonesia and Banco de México, February 16-17, 2008, 42-81.

BIKKER, J.A. and SPIERDIJK, L. (2009b), *The impact of competition and concentration on bank solvency*, mimeo.

[33] DNB Working Papers: www.dnb.nl/dnb/home/nieuws_en_publicaties/publicaties_dnb/dnb_working_papers_reeks/dnb_working_papers/nl/46-148637.html.

[34] Tjalling C. Koopmans Research Institute Discussion Papers: www.uu.nl/uupublish/tjallingkoopmans /publications/discussionpapers/discussion papers/45511main.html.

BOONE, J. (2001), "Intensity of competition and the incentive to innovate", *International Journal of Industrial Organization* 19, 705-726.

BOONE, J. (2008), "A new way to measure competition", *Economic Journal* 118, 1245-1261.

BOS, J.W.B., 2004, "Does market power affect performance in the Dutch banking market? A comparison of reduced form market structure models", *De Economist* 152, 491-512.

DE LANGE VAN BERGEN, M. (2006), *The determinants of banking competition: a world-wide Lerner index approach*, Doctoraalscriptie Rijksuniversiteit van Groningen.

DEMSETZ, H. (1973), "Industry structure, market rivalry and public policy", *Journal of Law and Economics* 16, 1-9.

HICKS, J. (1935), "Annual survey of economic theory: monopoly", *Econometrica* 3, 1-20.

KOETTER, M., KOLARI, J.W. and SPIERDIJK, L., *Testing the 'quiet life' of U.S. banks with Lerner indices*, Universiteit van Groningen.

MARCUS, A. J. (1984), "Deregulation and bank financial policy", *Journal of Banking & Finance* 8, 557-565.

OECD (2000, 2002, 2004), *Bank Profitability; Financial Statements of Banks*, OECD, Paris.

PANZAR, J. and ROSSE, J. (1987), "Testing for 'monopoly' equilibrium", *Journal of Industrial Economics* 35, 443-456.

SCHAECK, K., CIHAK, M. and WOLFE, S. (2006), *Competition, concentration and bank soundness: New evidence from the micro-level*, IMF Working Paper WP/06/185, Washington, D.C., International Monetary Fund.

SCHAECK, K. and CIHAK, M. (2007), *Banking competition and capital ratio*, IMF Working Paper WP/07/216, Washington, D.C., International Monetary Fund.

VAN LEUVENSTEIJN, M., BIKKER, J.A., VAN RIXTEL, A.A.R.J.M. and KOK SØRENSEN, C. (2007), *A new approach to measuring competition in the loan markets of the euro area*, DNB Working Paper No. 143/ECB Working Paper No. 768.

VAN LEUVENSTEIJN, M. KOK SØRENSEN, C., BIKKER, J.A. and VAN RIXTEL, A.A.R.J.M. (2008), *Impact of bank competition on the interest rate pass-through in the euro area*, DNB Working Paper No. 171/ECB Working Paper No. 885/Tjalling C. Koopmans Research Institute Discussion Paper Series No. 08-08, Universiteit Utrecht.

4.8. APPENDIX

Table A.1: Correlation Coefficients Between Indicators and the Index (46 Countries, 1996-2005)

	Boone	Panzar-Rosse	SCP	Cournot	Profit eff.y	Cost-eff.y	RoC	RoA	C/I	TC/TI	NIM	Cost-margin	No. of banks	C5	Index
	neg[a]	*pos*	*Neg*	*neg*	*amb (p)*[b]	*pos*	*neg*	*neg*	*amb (n)*	*amb (n)*	*neg*	*neg*	*pos*	*neg*	
Boone	1.00	-0.34 **	-0.20	-0.13	0.36	0.11	0.18	0.06	0.07	-0.23	-0.21	0.00	0.06	0.11	-0.14
P-R		1.00	-0.04	-0.03	-0.03	0.09	-0.17	-0.28 *	0.02	0.17	-0.02	-0.22	0.09	0.03	0.33 **
SCP			1.00	0.29 **	-0.07	-0.05	0.12	0.02	-0.27 *	-0.08	0.07	-0.09	-0.02	-0.15	-0.05
Cournot				1.00	-0.12	-0.26 *	0.25 *	0.17	-0.21	-0.06	0.20	0.21	-0.31 **	0.35 **	-0.42 ***
P. eff.y					1.00	0.48 ***	0.33 **	0.10	-0.46 ***	-0.38 **	-0.23	-0.16	0.24	0.06	0.37 **
C. eff.y						1.00	-0.02	-0.08	-0.36 **	-0.36 **	-0.25 *	-0.25 *	0.32 **	0.05	0.53 ***
RoC[1]							1.00	0.73 ***	-0.39 ***	-0.34 **	0.20	0.18	-0.28 *	0.30 **	-0.30 **
RoA								1.00	-0.03	-0.34 **	0.57 ***	0.59 ***	-0.26 *	0.21	-0.50 ***
C/I									1.00	0.37 **	0.19	0.42 ***	0.05	-0.08	-0.42 ***
TC/TI										1.00	0.11	0.06	-0.05	-0.14	-0.20
NIM											1.00	0.62 ***	-0.21	0.03	-0.63 ***
CM												1.00	-0.18	0.00	-0.58 ***
# Banks													1.00	-0.55 ***	0.51 ***
C_5														1.00	-0.37 **

a. Correlation between the Boone indicator and competition is negative, et cetera.
b. Correlation between profit X-efficiency and competition is theoretically ambivalent, but turns out positive in practice (p)

Explanation: Asterisks indicate significance levels: 1, 2 or 3 asterisks indicate the 90%, 95% or 99% confidence levels, respectively. Shading of the correlation coefficients indicates where negative correlation is expected. (For the ambivalent Profit efficiency variable, this was done 'in retrospect'). The names of variables included in the Index are printed in boldface in the first and last columns.