Journal of Money and Economy
Vol. 11, No. 4, Fall 2016
pp. 443-466

Relationship between Structure and Performance in the Banking Industry of Iran

Akbar Najahi* **Yeghaneh Mousavi Jahromi**[†]
Farhad Khodadadkashi[‡] **Golam Ali Haji**[§]

Received: 25 Jul 2018 Approved: 4 Sep 2018

The purpose of this study is to investigate the causal relationship between the structure and performance in the banking industry of Iran. In doing so, the data and information of public and private banks from 1996 to 2015 is examined using Toda-Yamamoto causality test (TY) and Autoregressive Distributed Lag (ARDL) approach. Herfindahl-Hirschman index (HHI) (the degree of market concentration) trend shows that over the last two decades, the degree of concentration has reduced. In other words, the structure of banking industry has moved away from monopoly conditions towards competitive markets. The degree of profitability and the ratio of net return on assets (ROA) are used for the performance of banking industry. According to the results, the banks' profitability rate has been increasing over the years under study. The results of Toda-Yamamoto causality test (TY) and Autoregressive Distributed Lag (ARDL) approach also show that in the banking industry of Iran, both in the short term and long term, there is a one-way causal relationship from performance to structure. In other words, the structuralist theory of the banking industry of Iran is not approved and the Chicago School theory of the causal relationship between performance and structure cannot be rejected.

Keywords: Market Structure, HHI, Performance, ARDL, Toda-Yamamoto Causality Test
JEL Classification: D22, G21, L13, L16, L22

1 Introduction

Banking industry is one of the accelerating engines in every economic system and it is important in policy making for various economic purposes. This industry affects both monetary and fiscal policies of the state directly and indirectly. Therefore, recognizing the organization of this industry and its

* Department of Economics and Management, Islamic Azad University of Arak, Iran; najahi.akbar@gmail.com (Corresponding Author)
† Department of Economics, Payam-e Noor University, Iran; yeganehmj@gmail.com
‡ Department of Economics, Payam-e Noor University, Iran; khodadad@pnu.ac.ir
§ Department of Economics and Management, Islamic Azad University, Arak, Iran; g-haji@iau-arak.ac.ir

structure-performance relationship can lead to optimal decisions in adopting monetary and fiscal policies.

In the last two decades, and especially with the emergence of private banks, structural changes have occurred in the banking industry of Iran, which resulted the industry to evolve from monopoly and oligopoly markets towards monopolistic competition and competitive market. In recent years, this structural change has been intensified with the implementation of Article 44 of the Constitution of the Islamic Republic of Iran (see Algar, 2015). According to structuralists' studies, structural changes can affect the performance of the industry and each active firm in a market. Studying the relationship between structural changes and market performance, affects the implementation of various policies, such as monetary policies at macro level and firm policies at micro level which results in optimal decisions.

Various studies have been conducted on the relationship between the structure and performance of an industry in different countries. These studies show a causal relationship between market structure and market performance. However, this causal relationship between the structure and performance of the industry has different degrees in different markets and the causality direction between these two variables, is not fixed. In other words, in some markets, the causality direction is from structure to performance, and in other markets it is from performance to market.

Different schools of economics have different proposals on the relationship of market elements (structure and performance). Structuralists believe that the direction of causality is from structure to performance; in contrast, the Chicago School advocates consider causality from performance to structure (Khodadadkashi, 2006).

This paper is an empirical analysis of banking market performance in Iran. The causal relationship between the structure and performance in banking industry is examined. In fact, it answers the question of whether there is a causal relationship between the market structure and the performance of the banking industry in Iran. Moreover, if this causal relationship exists, what is the direction?

In this study, the relationship between the degree of market concentration and profitability rate has been analyzed. For this purpose, data of the public and private banks were utilized from 1996 to 2015 using time series analysis, Toda-Yamamoto causality test (TY) and Autoregressive Distributed Lag (ARDL) approach. Therefore, the research method is descriptive-causal to answer the questions.

First the literature on the topic being studied. Then the theoretical foundations and the model and indices of the measurement of the variables being introduced. In the third section, the model variables are estimated. In the final part, the results and recommendations are presented.

2 Literature Review

Economic theories predict that monopoly will lead to higher prices and bigger performance losses compared to a competitive environment. This theory predicts that the degree of monopoly and the scale of the banking industry will affect its performance and vice versa. In other words, as performance will affect conduct and structure, the structure will also affect conduct and performance of firms (Matthews et al., 2014, 141). This model, which is known as Structure–Conduct–Performance (SCP) Paradigm, can be summarized as relation (1):

$$Performance \rightarrow conduct \rightarrow Structure \quad (1)$$

There are different views on the direction of the causality between market structure and performance. According to SCP School, the main points studied by different economists are the structure, conduct, the internal organization of the enterprises active in the industry and the external conditions governing the industry. The study of the effect of firm's structure on the performance initiated with the application of the market power and optimal structure theory. However, the theory of market power is another expression of the SCP theory. This approach is largely based on the results of Bain (1951, 1956), who believes that firms have a higher profitability rate in industries with a high degree of concentration.

According to this approach, market performance is strongly affected by the conduct of firms as well as market structure. Market structure is largely affected by the relative size of firms and the conditions governing the market for the entry of new firms or the possibility of easy and low-cost exit of firms and the market demand capacity. According to the structuralists' approach, the causality direction between elements of the market is from structure to conduct and then performance. Nevertheless, many economists have criticized this approach so far (Cowling & Waterson, 1976). Some economists with the introduction of the Chicago School try to introduce monopoly as a short-term phenomenon. The group argue that in the long-term markets are in a competitive environment. Thus, indices such as concentration are not important in this regard.

Another group of critics is the Austrian school advocates who are like the Chicago school in their principles, except that they do not consider monopoly as a short-term and insignificant phenomenon (Baumol, 1988). They believe that monopoly is a reality that can continue in the long-term. According to this school, firms can benefit from a monopoly situation with cost efficiency (Cowling & Waterson, 1976; Baumol, 1988).

Market entry qualifications or barriers to entry, are other structural variables shaping the conduct and performance of firms. As the barriers for market entry are fewer, the firms are more likely to be competitive (conduct), and because of the threat of new entrants, they have to comply with the pricing rule based on the marginal cost (performance). Structuralists' view, or in other words, SCP paradigm can be represented in the form of the following equation:

$$Performance = f(external\ conditions\ and\ the\ internal\ organization;\ conduct;\ structure) \quad (2)$$

In most of the empirical studies, the profitability rate is used as a performance variable for examining the effects of the structure on performance. Market share, industry concentration index, barriers to entry into the industry are examined as structural variables.

In contrast to the Chicago School, UCL School is formed, which criticizes the SCP paradigm. Causality is not considered from structure to performance in this approach. Unlike SCP paradigm, this school does not consider industrial concentration as a factor in the formation of monopoly power. The principles of this school are formed by famous economists who mainly worked at both the University of Chicago and University College London, and are introduced by economists such as Stigler (1971), Brozen & Bittlingmayer (1982), McGee (1971), Alchian & Demsetz (1972), Posner (2014) and other writers who opposed structuralists' ideas in the 1970s. The advocates of this school believed that monopoly was a short-term phenomenon, and that in the long-term, rival firms would eliminate monopoly power. Thus, in the long-term, monopoly is considered to have less importance. According to the School's view, the monopolies seen in the real world are rooted in government privileges and governing institutions that create monopoly conditions for a firm.

Alchian & Demsetz (1972) attributes the cause of monopoly to the superior performance of monopoly firms. The effect of the three elements of the market on each other and their relationship are explained in the following equation (Sadraei Javaheri, 2011):

$$Structure = f(external\ conditions\ and\ internal\ organization;\ conduct;\ performance) \quad (3)$$

The two structuralist schools of Harvard and Chicago are different in various aspects. In their analyses, in addition to pure economic theories, structuralists give importance to the empirical studies, whereas Chicago School scholars rely heavily on pure economic theories and, especially, on competition theory.

In studying the relationship between the three elements of the market, one needs to study the measurement methods of each of these elements. The market structure, or in other words, the way in which production units are organized in a particular industry is among the topics covering a wide range of industry conditions (from monopoly to competition). In fact, one can find the degree to which the competition in each market is limited by identifying the structure of an industry (such as the banking industry) and calculating the intensity of the concentration.

The intensity of the concentration is one of the concepts discussed in the economic texts with many views regarding its measurement. In a classification of the concentration indices, they are divided into two categories of absolute concentration and dispersion indices. Indices of absolute concentration focus on markets division between firms, whereas dispersion indices focus on the degree of dispersion of the market share of firms in an industry. Among the indices of absolute concentration, one can refer to K Firm Concentration Ratio, HHI, Hannah-Kay Indices, and entropy index (Bajo & Salas, 2002). It should be noted that most absolute concentration indices are also affected by the change in the dispersion of the size of the market or the size of the corporate market (Sadraei Javaheri, 2011, 107).

SCP studies in United States have significantly used the concentration of the deposit market to measure market structure in the banking industry. The same studies in Europe have used the total asset index. However, it should be noted that the concentration of deposits is not necessarily the best index for competition, because the leading banks may have a high competitive market. However, due to the simplicity and the limitations of the required data, this ratio is one of the most commonly used indices in empirical definitions, which accumulates the market share of large banks (usually 3, 5 or 10 banks) in the economy (Khodadad Kashi, 2012).

3 Empirical Studies

In determining the structure of an industry, two types of models are used: the structural and non-structural models. Structural models are based on two

theories: Structure–Conduct–Performance (SCP) and Efficiency Structures (ES). In the structural models, calculating indices such as concentration, barriers to entry and product differentiation using market share determines the market structure. In contrast, non-structural models are based on the theoretical foundations of microeconomics. Non-structural models include Lerner, U Davis, and Panzer-Ross models (Ramzi et al., 2014). The empirical studies in Iran are mainly based on structural models using indices such as concentration and entry barriers, and non-structural models are less studied.

Early surveys of the literature include Gilbert (1984). It is probably the most comprehensive of these. He reviews early studies of US banks and finds that 32 out of the 44 studies support the SCP paradigm. Studies using European data also find support for the SCP paradigm. Rotella et al. (2004) offers a more recent review. This review focuses more on the new industrial organization empirical studies. However, it provides a relatively comprehensive review of the structure-conduct- performance from both a theoretical and empirical perspective. (Simatele, Mishi & Ngonyama, 2018).

Gavurova, et al. (2017) investigate the relationship between structure and performance in the banking industry at the European Union (EU) between 2008 and 2015. The researchers test the presence of SCP paradigm in the EU. The presence of this paradigm was verified using the Granger causality test for panel data. The results of analysis show that under the studied conditions only the one-way relationship running from banking sector performance to banking market concentration is approved. The findings do not confirm the presence of the SCP paradigm, but are in line with the "quiet life hypothesis" (QLH), thus indicating there is a negative relationship between concentration and performance at European banking market.

Tarus and Kimeli Cheruiyot (2015) have studied the relationship between structure and performance in the banking industry in Kenya based on structural models. The researchers considered the efficiency of the industry as the market structure, and calculated HHI for the degree of market concentration. They considered the information of 44 commercial banks from 2000 to 2009 to examine the relationship between structure and performance in the banking industry in Kenya. They introduce a generalized least squares (GLS) method to estimate their model. They show a significant relationship between performance of banks and efficiency in the banking industry of Kenya.

Ye et al. (2012) study the banking industry in China. Based on the structural models and SCP paradigm, the researchers have presented five hypotheses to study the relationship between structure and profitability. Using

the data of 14 large banks in China from 1998 to 2007 and panel data method, Ye et al. conclude that the Chinese banking industry is changing from the monopolistic market to a competitive market. By estimating their model, they show that none of the two hypotheses - SCP and performance structure - is supported in Chinese banking market and the only hypothesis supported by China's banking industry is the “relative strength” theory of the market.

Scholtens (2000) have studied the relationship between competition, growth and performance in the banking industry of 7 industrial countries using structural models. Scholtens (2000) main hypothesis in this study is the profit growth of banks that depends on their size. In this study, considering the financial data of 100 banks in 7 industrial countries, Scholtens (2000) shows that there is a weak relationship between the profitability of banks and banking concentration. Moreover, he shows that banking concentration does not lead to a monopoly profit in the banking sector of these countries. Brett's studies also confirm the relationship between the bank's internal characteristics and earnings growth. Based on Brett's study, profit growth (performance) is positively correlated with the size of a bank's capital. He also shows that the relationship between the profit and size with the bank's assets is much stronger.

Also, for Iran’s economy, Shahiki Tash et al. (2015) have used structural models and written a paper entitled “Investigating the relationship between market structure and profitability coefficient in the banking industry of Iran.” The purpose of this study is to examine the relationship between structure and performance. The study considers market concentration as a structural variable and profitability as a performance variable. They estimate firms' profits using demand function and show that market power of firms is a performance of the market share of firms. In doing so, financial information of 17 private and public banks are considered for 2008-2012 and the model is estimated based on panel data method. According to the results of this study, the coefficient of concentration index in the deposit market is statistically positive and insignificant. In other words, the concentration on the deposit market of Iran's banking industry does not lead to profit making, but as the market share of the banks becomes higher in the asset dimension, the profitability of the bank increases positively. In this study, by mentioning the structural changes in the 2000s, some hypotheses have been tested. According to the results of hypothesis testing, the first hypothesis stating the relationship between market concentration coefficient and profitability in Iran's banking industry was rejected with 99% confidence. In other words, based on four years data, 2008-2012, there is no significant relationship between the

structure and the performance. The second hypothesis concerning the comparison of the profitability ratio between private and public banks is not rejected and the difference between the average profitability of the two statistical societies is significant.

In a study, Razmi et al. (2014) explores the structure of the banking industry in Iran. He uses both structural and non-structural models to determine the structure of the market. With the introduction of structural and non-structural models, the study identified the industry concentration using HHIs, the concentration of superior banks, entropy, comprehensive industrial concentration, Hall-Tideman and Ross Blooth during 2010-2013 and shows that, despite the concentration of the banking industry in Iran, the downward trend of concentration and the level of competition during the study period are incremental. In this study, the empirical model of Panzer-Ross, which is a non-structural model, is estimated using the panel regression method. With various specifications of this model, it has been shown that the banking industry of Iran has been in a state of monopoly. In this study, Honarvar refers to the existence of observations concerning the negative relationship between profitability and size of the bank. He uses quantum regression method to examine the relationship between bank size and performance of bank. The results of this method also show no clear relationship between size and performance in the banking industry in Iran.

Khodadadkashi and Jafari Lilab (2012) have used structural models to examine the relationship between the structure and performance of the banking industry based on structuralists' views. In this study, productivity is considered as the performance index, and Tornqvist-Theil index (TTI) is introduced to measure the productivity of banks. This index is calculated considering the data of 11 banks. The researchers believe that the internal characteristics of the bank affect its productivity, so the factors affecting this variable are studied and evaluated as well. One of the factors mentioned in the model as an explanatory variable is the bank size. The indices of concentration, ownership and market share of banks are used to calculate market structure. HHI of concentration, which is a structural model, is used to examine the effect of concentration on the performance of the banking firm. According to the results of this study, among the factors affecting the productivity of banks, bank capital is more effective than other variables. Although the concentration index, as an explanatory variable of market structure, has a negative and significant effect on productivity (as an explanatory variable of performance), the researchers do not confirm structuralists' theory in the banking industry of Iran.

4 Data Descriptions

Economists have much attempted to understand the relationship between structural and functional elements of the market and in many empirical works they have sought to understand the relationship between the profit rate as an indicator for the performance and market concentration as the structural variable. Before Cowling and Waterson (1976) model presentation, none of these studies had strong discursive bases and the studies were more based on Adhoc models and statistical analysis in which researchers discarded some of the important variables. Cowling and Waterson (1976) show that in the empirical studies the relationship between performance and structure can be directly examined, like our empirical model that only performance and concentration variables have been investigated.

One of the most important structural variables mentioned in most studies is concentration. The most common measure of concentration which is the only criterion used by financial institutions is HHI. It is defined as the sum of the square of the market share of banks in the market. This index is calculated based on the following equation:

$$\text{HHI} = \sum_{i=1}^{n} S_i^2 \tag{4}$$

In this equation, S_i is the market share of the i-th firm (bank) and n is the total number of firms (banks) in the industry.

The minimum value of this index is obtained when all firms have a market share equal to each other. In this situation, the distribution of the market share of the firm is zero and the rate of this index is 1 / n. In markets where their structure is in competition, the index will be close to zero as the number of firms in these markets is very high. Therefore, the value of 1/n tends towards zero. For the markets where their structure is near monopoly, this index tends towards one and in absolute monopoly, the index is one. Therefore, HHI always has a number between 1/n and one (Matthews et al., 2014, 185).

$$1/n \leq HHI \leq 1 \tag{5}$$

The advantage of this index is the full use of the distribution of the market position of banks. According to its calculation method, this method gives the largest value to large banks and possible values for 1/n is one. This index reaches its minimum when the share of competitors equals and reaches the highest level of net monopoly. We can use the percentage of capitals held by 5 and 10 largest banks and the total squares of assets held by each bank related to HHI as an alternative to this variable (Khodadadkashi, 2012, 189).

As it has already been mentioned, the purpose of this study is to examine the interaction between structure and performance in the banking industry of Iran. The market structure in this study is calculated using HHI. with the information and financial statements of 30 public and private banks, this index is calculated for 1996-2015. For calculation of this index, each year, the share of banks' assets is considered as the total assets of the banking industry and is calculated using equation (4).

Table 1
The Degree of Market Concentration (HHI) in the Banking Industry of Iran

Year	***1996***	***1997***	***1998***	***1999***	***2000***	***2001***	***2002***	***2003***	***2004***	***2005***
Index	0.159	0.175	0.172	0.169	0.165	0.158	0.157	0.145	0.138	0.124
Year	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
Index	0.116	0.109	0.103	0.102	0.097	0.092	0.085	0.073	0.073	0.070

Source: Financial statements of banks (1996-2015)

In Table 1, HHI, which represents the degree of market concentration from 1996 to 2015, has always been in a downward spiral and the market structure has distanced from monopolistic markets. In other words, during the last two decades, the structure of the banking industry has been moving away from monopolistic markets to competitive markets with the entry of private banks.

Overall, the empirical studies have used the relationship between performance-structure by using two general approaches for measuring bank performance. One approach uses the price of products or services and the other approach uses profitability index. Studies using the price-driven index often use average loan rates, deposit rates, or revenues from fees. Using such variables has been criticized for several reasons: average interest rates calculated from the balance sheet (denominator) and earnings (nominator) are reserved variables (loan portfolios at the end of the period under review) that are combined with flow variables (profit earnings in the studied period) (Rezaei, 2014). Another commonly accepted method for measuring bank's performance is the use of profitability index (ROA and ROE)[1]. The most important advantage of them is their simplicity and the fact that only a single number can represent the performance of banks as multi-product firms.

ROA is one of the important indices in determining the bank's ability to manage the optimum use of the real capital and financial resources in creating

[1] Average Return on Assets (Net Income/Total Assets) and Average Return on Assets (Net Income/Total Equity).

profits. The low value of this ratio shows that the bank has not used its assets well. This ratio is calculated by dividing net income into its assets[1] (Pajouyan, 2008, 93).

ROA method, which actually represents the degree of profitability in the banking industry, is used to obtain the performance.

$$\text{ROA} = \frac{E}{P}.\frac{P}{A} = \frac{E}{A} \tag{6}$$

Here, E is the net profit, P is the sales volume and A is the total assets. Ultimately, this ratio is calculated by dividing the net interest received on assets. In Iran, instead of net gain, the share of the bank is derived from the combined revenues and income earnings. In other words, the proportion of the total revenues of all banks and their total assets is obtained. To calculate this index, the information and financial statements of 30 public and private banks for years 1996-2015 are considered.

Table 2
Profitability (ROA) of the Banking Industry of Iran

Year	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
ROA	0.116	0.106	0.155	0.126	0.177	0.295	0.285	0.459	0.552	0.263
Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
ROA	0.373	0.277	0.403	0.413	0.501	0.631	0.595	0.579	0.251	0.164

Source: Financial statements of banks (1996-2015)

The results of the calculations according to Table 2 show that the performance of the banking industry in Iran has improved during the study period and the degree of profitability of the industry (net profit to asset) has experienced an increasing trend. This means that with the entry of private banks into this industry, not only the degree of profitability of the industry has not been reduced, but also this index has improved year by year. The activity of private banks has gradually happened over the past two decades, and several new banks have entered the market each year. This shows that with the increase in the number of private banks, the degree of profitability of the industry has increased, which could be from the improvements in the operational processes of large banks due to arrival of new banks that has forced them to modify their processes.

[1] Due to using and implementation of non-usury banking in Iranian banks, the removal of interest and the replacement of earnings in its place and non-separation of profits in the financial statements of banks, ROA is calculated by dividing the bank's share of total income and income from equity into each bank's assets.

5 Methodology

5.1 Toda-Yamamoto Causality Test (TY)

Toda-Yamamoto (1995) suggest a simple method as estimating an adjusted VAR model to study the Granger causality. They argue that this method is valid even in the presence of a co-integration relationship between variables. In this method, first, the number of optimal lags (k) of VAR model, then the maximum degree of Co-integration (dmax) form a VAR model with the number of lags (k + dmax). However, the interruption selection process will be valid at k≥dmax, so if we consider the following two-variable model, then the Toda-Yamamoto causality test can be determined as follows:

$$Y_t = \alpha_0 + \beta_{1i}\sum_{i=1}^{k} Y_{t-i} + \beta_{2i}\sum_{j=k+1}^{d\max} Y_{t-j} + \gamma_{1i}\sum_{i=1}^{k} X_{t-i} + \gamma_{2i}\sum_{j=k+1}^{d\max} X_{t-j} + \varepsilon_{1t} \quad (7)$$

$$X_t = \alpha_1 + \lambda_{1i}\sum_{i=1}^{k} X_{t-i} + \lambda_{2i}\sum_{j=k+1}^{d\max} X_{t-j} + \delta_{1i}\sum_{i=1}^{k} Y_{t-i} + \delta_{2i}\sum_{j=k+1}^{d\max} Y_{t-j} + \varepsilon_{2t} \quad (8)$$

The test statistic used is Wald test that has the asymptotic χ^2 distribution with a degree of freedom equal to the number of zero limits.[1] Zapata and Rambaldi (1997) state that the advantage of this method is we do not need to know the co-integration properties of the system and only the knowledge of VAR model rating and the degree of the maximum Co-integration of the variables are sufficient for the test (Rezaei, 2014). For doing the conventional Granger causality test, it is necessary to study the long-term and co-integration relation between the variables, whereas in Toda-Yamamoto causality, information on the co-integration properties of the system is unnecessary.

5.2 ARDL Approach

Many studies have used Johansen's approach to examine the long-term relationship between variables. However, in recent years, Pesaran et al. (2001) have introduced Autoregressive Distributed Lag (ARDL) approach, which has some advantages compared to Johansen's technique and to Vector Autoregressive (VAR) and Vector Error Correction (VECM) models. Unlike Johansen's approach, there is no need to know the co-integration degree of the variables in the model and the number of co-integration vectors is determined. This approach has the following advantages over VECM method.

[1] Attention should be paid that Wald limitation test is done only on the main lags(k).

First, ARDL approach is suitable for smaller samples, whereas larger samples need to rely on the results of Johansen's Approach (VAR and VECM models).

Other co-integration methods (including Johansen) need the same degree of integration of variables, and if only one of the variables is stationary with one degree of difference, the first-order difference of all variables (even the variables at stationary level) has to be used. This ends in the loss of a large volume of information in stable variables. However, ARDL approach can be used for variables with varying degrees of integration.

In ARDL approach, it is possible to consider different optimal lags of each variable at different stages of the estimation, whereas this is impossible in Johansen's Approach.

Estimates of ARDL approach are unbiased, effective given the avoidance of problems such as auto-correlation, and endogeneity. In addition, this method simultaneously estimates long-term and short-term relationships between the dependent variable and other explanatory variables of the model.

Bound test in ARDL is a new method for determining the long-run relationship between a dependent variable and a number of estimators. Y_t and the two types of equation are estimated to examine the co-integration relationship between the variables if X_t as dependent variable.

$$\Delta X_t = a_1 + \sum_{i=1}^{k} b_{i1}\Delta X_{t-j} + \sum_{i=1}^{k} c_{i1}\Delta Y_{t-j} + \sigma_1 X_{t-1} + \sigma_2 Y_{t-1} + \varepsilon_{1t} \quad (9)$$

$$\Delta Y_t = a_2 + \sum_{i=1}^{k} b_{i2}\Delta X_{t-j} + \sum_{i=1}^{k} c_{i2}\Delta Y_{t-j} + \omega_1 X_{t-1} + \omega_2 Y_{t-1} + \varepsilon_{2t} \quad (10)$$

In equation 9, Δ is the differential operator, X is the dependent variable and Y is the vector of independent variables. In equation (10), Y is the dependent variable, and X is the vector of independent variables, ε is the error term, t shows the time and k is the number of optimal lags that can be estimated using Akaike Criterion (AIC), Schwartz-Bayesian (SBC), (AIC) or Hanan-Quinn (HQC), or modified R^2 to see whether ω, σ, a, b coefficients are measurable parameters.

In equation (9), where X is the dependent variable, we test the null hypothesis of the absence of a long-run relationship between variables (H_0: $\sigma_1 = \sigma_2 = 0$) against the alternative hypothesis (H1: $\sigma_1 \neq \sigma_2 \neq 0$) using F statistics. However, the distribution of this F statistic is not standard regardless of I(0) or I(1) of the independent variables. In doing so, Pesaran et al. (2001) propose

the appropriate critical values considering the number of independent variables and the existence or absence of cross-section or time trend in them. These statistics include two sets: one set is estimated assuming that all variables are I (0) and the other set estimated based on all variables as I (1). If the calculated F statistic exceeds the limit of the critical values provided by Pesaran et al., we reject the null hypothesis of no long-term relationship between the variables. Thus, one can deduce a one-way Granger-causality relation from X_t to Y_t. If the calculated F statistic is less than the lower limit of this range, then the null hypothesis cannot be rejected, and there is no Granger-causality relationship between the variables. If the calculated F statistics is within the range of critical values, then the result cannot be determined or inferred. In this case, this procedure should be repeated.

6 Results

6.1 Stationarity Test of the Variables

Before model estimation, we need to investigate the stationarity of the model variables. Also the ARDL bounds test is based on the assumption that the variables are I(0) or I(1). So, before applying this test, we determine the order of integration of all variables using the unit root tests. The objective is to ensure that the variables are not I(2) so as to avoid spurious results.(Pesaran et al. , 2001)

One of the most common unit root tests for time series variables is Augmented Dickey Fuller test, which is used in this study. In this test, if the calculated value of the statistic is less than the critical value, then the null hypothesis of the existence of unit root in the variable is not rejected.

Table 3
The Results for Unit Root Test of Augmented Dickey Fuller for the Variables

Variable	T value	Critical Values for the Dickey-Fuller			Result
		1%	5%	10%	
H	0.59	-3.815	-3.029	-2.65	non-stationary
P	-1.85	-3.835	-3.029	-2.660	non-stationary
ΔH	-6.231	-3.875	-3.040	-2.660	stationary
ΔP	-4.23	-3.857	-3.04	-2.66	stationary

Source: Research Findings

The results of the stationarity tests show that both variables are non-stationary. These results are given in Table 3. The ADF test applied to the first

difference of the data series reject the null hypothesis of nonstationary for both variables.

6.2 Toda-Yamamoto (TY) Causality Test

It is necessary to study the long-term and co-integration relationship between variables to conduct the conventional Granger causality test, whereas in TY causality, information about the coherent co-integration of the system is not necessary. As mentioned, in this method, first, the number of optimal lags (k) of VAR model, then the maximum co-maximal value (dmax) should be determined. Moreover, a VAR model with (k + dmax) number of lags should be formed. According to the results from augmented Dickey–Fuller unit root test, the degree of reliability is one, and based on AIC criteria, Schwartz and final predictive error; the optimal interruption is two.

Table 4
Selection of the Optimal Interruption

Lags	Log L	AIC	SC	HQ
1	85	-8.54	-8.24	-9.34
2	91	-9.56	-9.09*	-9.5*
3	96*	-9.7	-9.07	-9.69*

Note. * indicates lag order selected by the criterion. *Source*: Research Findings

Equations (11) and (12) with 3 intervals and the Wald test for coefficients obtained from (dmax+k=1+2=3) are used to investigate TY causality relationship between the performance and structure.

$$H_t = c_0 + \alpha_{1i} \sum_{i=1}^{3} H_{t-i} + \beta_{1j} \sum_{j=1}^{3} P_{t-j} + \varepsilon_{1t} \quad (11)$$

$$P_t == c_0 + \alpha_{2i} \sum_{i=1}^{3} H_{t-i} + \beta_{2j} \sum_{j=1}^{3} P_{t-j} + \varepsilon_{2t} \quad (12)$$

Here, H is the structure and P is the performance. The results of the estimation of the above equations are as follows.

Ht = 0.3 + 0.46 H (-1) + 0.56 H (-2) - 0.18 H (-3) – 0.013 P (-1) – 0.021 P (-2) – 0.006 P (-3)
(2.74) (1.53) (2.02) (1.46) (-2.02) (2.63) (-0.66)

Pt = 0.68 - 6.73 H (-1) + 3.75 H (-2) + 0.24 H (-3) + 0.6 P (-1) – 0.007 P (-2) – 0.68 P (-3)
(1.18) (-0.45) (0.25) (0.036) (1.8) (0.01) (-1.28)

The results of Wald test indicate the significance of the coefficients with the interruption of the variables used in equations (11) and (12). As shown in Table 9, there is a one-way causality relationship from performance to structure.

Table 5
Results of Toda-Yamamoto (TY) Causality Test

Dependent variable	Independent variable	Null hypothesis	Wald test (χ^2)	Prob	Test result
H	P	$\sum_{i=1}^{3}\beta_i = 0$	8.31	0.0039	The null hypothesis is rejected and performance is the granularity cause of structure
P	H	$\sum_{i=1}^{3}\alpha_i = 0$	0.79	0.37	The null hypothesis is not rejected and structure is not the granularity cause of performance

Source: Research Findings

6.3 ARDL Approach

After examining the stationary condition of the variables, the ARDL model is estimated using bounds cointegration test provided by Pesaran, Shin and Smith (2001). Using this method, the long-term equilibrium relationship between the model variables is tested. In this study, two types of relationships are examined. For the cointegration test, the following UECMs is estimated:

$$\Delta H_t = a_1 + \sum_{i=1}^{k} b_{i1}\Delta H_{t-j} + \sum_{i=1}^{k} c_{i1}\Delta P_{t-j} + \sigma_1 H_{t-1} + \sigma_2 P_{t-1} + \varepsilon_{1t} \quad (13)$$

$$\Delta P_t = a_2 + \sum_{i=1}^{k} b_{i2}\Delta H_{t-j} + \sum_{i=1}^{k} c_{i2}\Delta P_{t-j} + \omega_1 H_{t-1} + \omega_2 P_{t-1} + \varepsilon_{2t} \quad (14)$$

Here, ε shows the error term, and the coefficients ω, σ, a, b are the measurable parameters. Now, in the above equations, the null hypothesis is based on the absence of long-term relationship between variables against the

alternative hypothesis. We use F statistics. The results of equations (12) and (13) are shown in Table 4.

The results of the table show that when the dependent variable is (H), the calculated F statistic is higher than the maximum critical values at different levels. Thus, the null hypothesis of no long-term relationship can be rejected and the long-term relationship between variables at different levels of confidence is confirmed. Thus, there is a Granger causality relationship from performance to structure. For the case where the dependent variable of performance is (P), the F-computational model is lower than the lower bound of the critical values at lower levels. The null hypothesis is based on the absence of a long-term relationship. In other words, the structure variable is not a function of Granger causality performance, and there is a one-way relationship between structure and performance.

Table 6
Results of Bounds Test

H0	F-statistic	I(0) Bound			I(1) Bound			Test result
		2.5%	5%	10%	2.5%	5%	10%	
Performance is not the Granger causality of structure	33.1	4.18	3.62	3.02	4.79	4.16	3.51	The null hypothesis is rejected and performance is the Granger causality of structure
Structure is not the Granger causality of performance	0.1.1	4.18	3.62	3.02	4.79	4.16	3.51	The null hypothesis is not rejected and structure is not the Granger causality of performance

Source: Research Findings

In this section, based on the results of the Co-integration bounds test, the long-term and short-term coefficients of the equations with F statistic are higher than the upper bound, confirming the Co-integration hypothesis. Thus, as the long-term relationship from performance to structure is confirmed, only one equation is estimated using ARDL method, i.e., the model where the structural is introduced as the dependent and performance as the independent variables.

Table 7
The Results of Estimating Dynamic Coefficient of the Model (1 And 2): (The Dependent Variable H)

Variable	coefficient	Standard Error	t-Statistic	Prob
H(-1)	0.867	0.037	23.142	0.000
P	-0.009	0.005	-1.622	0.129
P(-1)	-0.003	0.007	-0.487	0.634
P(-2)	-0.023	0.008	-3.036	0.010
C	0.023	0.007	3.163	0.008
R-squared	Adjusted R-squared	F-statistic		
0.995	0.994	701.307		

Source: Research Findings

The results of Table 7 show that all coefficients have the expected sign except for the coefficient P (-1), and all the other coefficients are significant at different levels of confidence and the variables explain 99% of the variations of the dependent variable. In addition, according to F-statistic, the whole model is statistically significant.

Table 8 shows long-run coefficients estimated by structural vector autoregressive (SVAR) with extensive interruptions. The results show that all variables are significant at 95% confidence level and, in the long-term, the banking industry's performance has negative effect on concentration (structure). In other words, with improvement in the performance of the industry, it become closer to competitive conditions. According to the results, with 1% increase in the performance variable, the structure decreases by 0.26%.

Table 8
Estimated Long Run Coefficients Using the ARDL Approach

Variable	coefficient	Standard Error	t-Statistic	Prob
P	-0.264	0.042	-6.364	0.000
C	0.175	0.014	12.285	0.000

Source: Research results

The normality of the components of the disruption, consecutive correlation and the heterogeneity of the variance between the sentences in the self-explanatory pattern are examined to ensure the correct choice for the length of the interruption. For this purpose, Lagrange Multiplier (LM), Jarque-Bera

(JB) and AR Root are used. Based on the above tests, the correct number of lags determined based on Schwartz is confirmed. The results of these tests are presented in Table 9.

Table 9
Results of JB, LM and AR Root Tests

Row	Test	Null hypothesis	Statistic	Prob	Test result
1	JB	The disruption sentences are normal	0.39	0.82	The null hypothesis is not rejected and the disruption sentences are normal
2	LM	There is no auto-correlation in disruption sentences	4.4	0.35	The null hypothesis is not rejected and the disruption sentences have no auto-correlation
3	AR Root	All roots are inside a single circle.			The system is stable

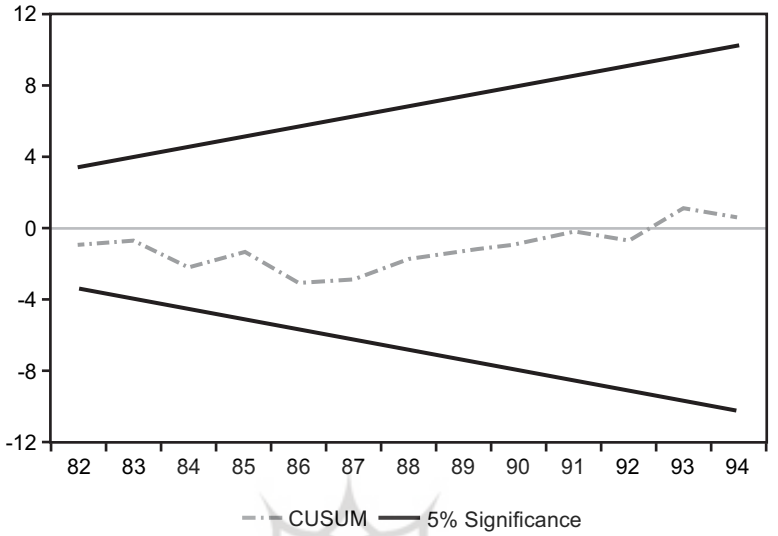
Source: Research results

6.4 Stability Test Results

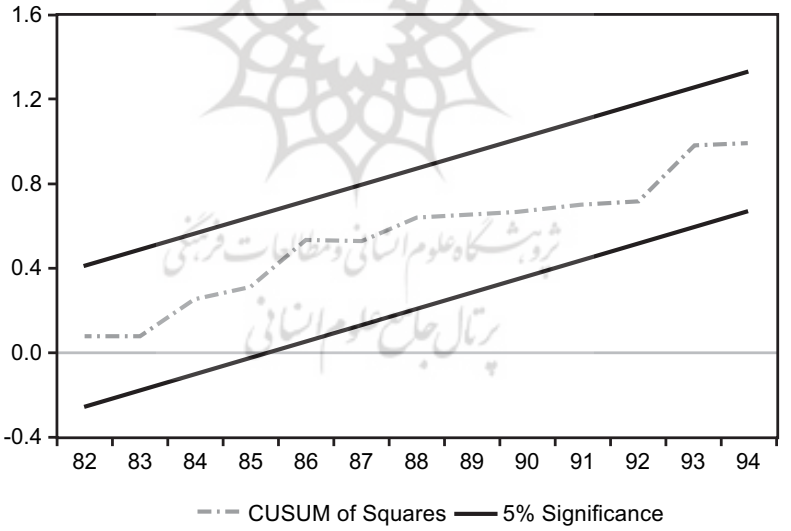
Cumulative Sum of Recursive Residuals (CUSUM) and Cumulative Sum of Square of Recursive Residuals (CUSUMSQ) have been used for the estimated ARDL model to ensure the stability of the estimated regression and the accuracy of the results obtained.

Pesaran and Shin (1997) propose these tests to determine the stability of both long-term and short-term parameters in the error correction model. However, Brown, Durbin and Ounce (1975) first proposed this test.

In these tests, the estimated values are plotted between two critical values at the 5% level, and if they do not exit these two edges, the null hypothesis, denoting the estimated regression being stable, cannot be rejected. The results of these two tests are shown in Figures 1 and 2. In both figures, the estimated values are between critical values and inside the two edges. Thus, the stability of the estimated regression cannot be rejected. Therefore, there is no structural failure in the estimated model.



Figures 1. Result of CUSUM test. *Source*: Research Findings



Figures 2. Result of CUSUMQ test. *Source*: Research Findings

7 Conclusion

The results show that over the past two decades, with the gradual entry of private banks, the structure of the banking industry has been moving from monopolistic markets to competitive markets. In other words, during the study period, from 1996 and 2015, HHI (the degree of market concentration) has always been declining, and the market structure has distanced from monopoly conditions. In addition, according to the results, one can state that the industrial banking industry is profitable, as the entry of private banks and their activities not only does not reduce the profitability of the banking industry, but also improve the profitability index (net profit to asset). Among the reasons is the improvement of operational processes in large and state-owned banks who forced to modify their processes with the arrival of new banks.

In examining the relationship between the two variables of the model, the structure and the performance, Toda-Yamamoto causality test (TY) and Autoregressive Distributed Lag (ARDL) approach are used, and a causal relationship is found between the performance and structure in the short-term and long-term, which is directed from the performance towards the structure of the industry. In other words, changes in the structure of the market during the study are due to changes in the performance of the industry, and with improved performance or an increase in the degree of profitability of the industry, the market structure has changed from monopolistic markets to competitive markets. The results of this study, like previous studies (Khodadadkashi and Jafari, 2012), show that the structuralists' theory in the banking industry in Iran is not confirmed, and the Chicago School theory of the existence of a causal relationship between performance and structure cannot be rejected.

The results of CUSUM and CUSUMSQ for ARDL model for the stability of the estimated regression also indicate no structural failure in the estimation model.

The results of diagnostics tests such as the normality of error components, serial correlation and heterogeneity of the variance among error sentences in the model are studied. In doing so, LM, JB and AR root tests are used. According to the results, the assumptions related to the error sentence are met and based in ROA analysis, in the two periods of 2005 and 2014, some shocks have been introduced into system from performance. These shocks are analyzed and interpreted as follows:

The first shock was from performance (profitability) in 2005 affecting the structure for several periods. The shock was formed by the imposition of small and medium-sized enterprises and the payment of loans to support the

purchase of housing for the deprived strata of the society. It reduced the speed of the banking industry moving from monopolistic markets to competitive markets. Since 2010, the process of moving towards competitive markets has started again. In other words, the intervention of the supervisory authority can prevent the banking industry from moving towards competitive markets.

The second shock was in 2014 due to various reasons. Considering the policy of previous years (2005-2013) in granting loans for various projects and the increase of time deposits of these loans in 2014, the volume of deferred bank claims, based on the Central Bank's data, increased to its highest level in the life of the banking industry of Iran. The volume of non-current facilities (including maturity, deferred and suspicious claims on banks) reached 815 billion IRR, which is 25% of the total banking facilities (this figure increased to 40% in 2016). The intensification of the economic sanctions in the four years leading to 2014 have led to a currency crisis, a decline in international banking relationships and high debt growth of the banks to the Central Bank. All this exacerbated the banks' credit crises, which finally ended in a decline in banks' income, which is reflected as decline in the number of facilities granted to deposits, high rates of profits, and price conflicts between banks and financial institutions and finally affected the performance and structure of Iran's banking industry. These shocks in performance lead to a reduction in the degree of profitability and in structure slows down the movement towards competitive markets. According to the results, one can predict that the effect of the second shock on the performance and structure variables will continue in the years 2016 and 2017.

Also, the profitability of the industry has increased from 1996 to 2013 and stopped since 2014. Regarding the confirmation of the causality relationship from performance towards structure, it is predicted that in 2017 with the decline in profitability, the banks collude with each other to prevent their decline in operational gains. One of the effects of this collusion is on the interest rates on deposits. These collisions could push the market structure from competitive to monopoly and oligopolistic markets.

References

Alchian, A. A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review*, 62(5), 777-795.

Algar, H. (2015). *Constitution of the Islamic Republic of Iran*. BookBaby.

Bain, J. S. (1951). Relation of Profit Rate to Industry Concentration. American Manufacturing, 1936–1940. *The Quarterly Journal of Economics*, 65(3), 293-324.

Bain, J. S. (1956). *Barriers to New Competition* [Vol. 3]. Cambridge, MA: Harvard University Press.

Bajo, O., & Salas, R. (2002). Inequality Foundations of Concentration Measures: An Application to the Hannah-Kay Indices. *Spanish Economic Review*, 4(4), 311-316.

Baumol, W. J. (1988). *Contestable Markets and the Theory of Industry Structure*. Harcourt College Publications.

Brozen, Y., & Bittlingmayer, G. (1982). *Concentration, Mergers, and Public Policy*. Free Press.

Cowling, K., & Waterson, M. (1976). Price-Cost Margins and Market Structure. *Economica*, 43(171), 267-274.

Durbin, J.; Brown, R.; Evans, J.; (1975) Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37 (2) pp. 149-192.

Gilbert, R. A., (1994). Bank Market Structure and Competition. *Journal of Money, Credit & Banking*, 16(4), 450 -617.

Gilbert, R., (1984). Bank Market Structure and Competition: A Survey. *Journal of Money, Credit & Banking*, 16(4), 450 -617.

Khodadadkashi, F. (2006), *Market Structure and Performance, Theory and Application in Iranian Industry* [Second Edition]. Tehran: Institute for Trade Studies and Research.

Khodadadkashi, F. (2012). *Industrial Economics (Theory and Application)*. Publication, 2013.

Khodadadkashi, F., & Jafari Lilab, P. (2012). Evaluation of Structural Elements on the Performance of Banking Industry in Iran. *Journal of Economic Research*, No. 98, spring 91, 90-69.

Matthews, K., & Thompson, J. (2014). *The Economics of Banking*. 3rd Edition, John Wiley & Sons Ltd, West Sussex PO19 8SQ, England.

McGee, T. G. (1971). *The Urbanization Process in the Third World*. London: G. Bell and Sons, Ltd.

Pajouyan, J., & Shafiei, A. (2008). Structural Analysis in the Banking Industry of Iran: Experimental Application of Davis U Index. *Quarterly Journal of Quantitative Economics* (Ex-Economic Reviews), Volume 5, Issue 4, winter 2008, 105-81.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds Testing Approaches to the Analysis of Level Relationship. *Journal of Applied Economics*, 16, 289-326.

Posner, R. A. (2014). *Economic Analysis of Law*. Wolters Kluwer Law & Business.

Razmi, M. J., Hossein, S., Mohammad, S., Aran, Z., Hossein, M., & Zarif Honarvar, A. (2014). *Study of the Firms Size Effect on Their Efficiency Based on DEA Approach* [Working Paper]. SSRN.

Rezaei, A. (2014). Toda-Yamamoto's Causality Test between Entrepreneurship and Unemployment Rate in the Industrial Sector: Case Study of Sistan and Baluchestan Province. *Quarterly Journal of Financial and Economic Policy*, Vol. 2, No. 7, Fall 2014

Rotella, J. J., Dinsmore, S. J., & Shaffer, T. L. (2004). Modeling Nest-Survival Data: A Comparison of Recently Developed Methods that Can be Implemented in MARK and SAS. *Animal Biodiversity and Conservation*, 27(1), 187-205.

Sadraei Javaheri, A. (2011). *Industrial Economics*. Industrial Management Organization, Tehran, 2011

Scholtens, B. (2000). *Competition, Growth, and Performance in the Banking Industry*. University of Groningen, Department of Finance, the Netherlands.

Shahiki Tash, M. N., Mahmoudpour, K., & Nayeni, M. (2015). The Relationship between Market Structure and Profitability in Iran's Banking industry (SCP Approach). *Trend, Quarterly Journal*, No.22, winter 2013, 37-13.

Simatele M., Mishi S., & Ngonyama N. (2018). Structure and Profitability in the Banking Sector. *Research Gate*.

Stigler, G. J. (1971). The Theory of Economic Regulation. The Bell Journal of Economics and Management Science, 43(2), 3-21.

Tarus, D., & Kimeli Cheruiyot, T. (2015). Market Structure-Performance Hypothesis in Kenyan Banking Industry. *International Journal of Emerging Markets* 09/2015, 10(4), 697-710. DOI: 10.1108/IJoEM-12-2012-0178.

The Central Bank of the Islamic Republic of Iran, the report of the performance of the banks of the country, a collection of volumes related to the years of 1999-2015.

Ye, Q., Xu, Z., & Fang, D. (2012). Market Structure, Performance, and Efficiency of the Chinese Banking Sector. *Economic Change and Restructuring*, 45(4), 337-358.

Zapata, H. O., & Rambaldi, A. N. (1997). Monte Carlo Evidence on Cointegration and Causation. *Oxford Bulletin of Economics and Statistics*, 59(2), 285-298.