Journal of Banking and Finance 87 (2018) 40–48

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Measuring banks' market power in the presence of economies of scale: A scale-corrected Lerner index

Laura Spierdijka*, Michalis Zaourasa

University of Groningen, Faculty of Economics and Business, Department of Economics, Econometrics and Finance, P.O. Box 800, 9700 AV Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 4 January 2017
Accepted 26 September 2017
Available online 28 September 2017

JEL classification:
D43
L13
G21

Keywords:
Market power
Lerner index
Economies of scale
U.S. commercial banks

ABSTRACT

A positive Lerner index indicates a welfare loss for consumers due to deviations from marginal-cost pricing. Such a welfare loss may not always be due to market power, though. In particular, marginal-cost pricing would result in negative profit for the firm in the presence of economies of scale. In such a scenario, a positive Lerner index could simply reflect the firm's need to earn non-negative profits rather than market power. We propose a novel, scale-corrected price-cost margin for firms that produce in the economies of scale range. We show that this measure is more informative about market power than the Lerner index itself. As an empirical illustration, we analyze market power in the U.S. banking sector using both the corrected and uncorrected Lerner index. The corrected Lerner index reveals significant market power for U.S. commercial banks during the 2000 – 2014 period.

1. Introduction

The degree of competition among the firms in a sector or industry has important welfare implications for both consumers and firms (see e.g. Lerner, 1934). This explains why there exists a vast theoretical and empirical literature about the measurement of competition. The theoretical contributions in this strand of literature often focus on the properties of new or existing measures of market power, while empirical studies typically apply one or more measures of market power to real-life data to assess the degree of competition in specific industries such as the banking sector (e.g, Perloff et al., 2007).

According to Blair and Sokol (2014, p. 325), *"the standard measure of market power, at least by economists, has come to be the Lerner index"*. The theoretical and historical foundations of the Lerner index have been extensively discussed in the literature (Amoroso, 1933; Lerner, 1934; Amoroso, 1938; 1954; Landes and Posner, 1981; Giocoli, 2012). A firm's Lerner index compares the market output price with the firm's marginal costs of production, where marginal-cost pricing is referred to as the *"social optimum that is reached in perfect competition"* (Lerner, 1934, p.168). A positive Lerner index is generally associated with the presence of market power. More precisely, however, the Lerner index measures the departure from the social optimum due to deviations from marginal-cost pricing, which some studies have referred to as the *'social loss due to market imperfections'* (Scitovsky, 1955; Fernández de Guevara and Maudos, 2004; Elzinga and Mills, 2011). Such market imperfections refer to circumstances causing the price to deviate from marginal cost, including – but not limited to – the presence of market power.[1]

Although the Lerner index has gained popularity in applied research, it has also been subject to criticism (e.g., Scitovsky, 1955; Landes and Posner, 1981; Cairns, 1995; Koetter et al., 2012). For example, Lindenberg and Ross (1981, p. 28) observe that the Lerner index *"does not recognize that some of the deviation of P from MC comes from either efficient use of scale or the need to cover fixed costs"*. Elzinga and Mills (2011) even call this *"the most important limitation of the Lerner index"*. This criticism of the Lerner index boils down to the observation that marginal-cost pricing would imply negative profits for the firm in the presence of economies of scale. In such a scenario, it is therefore unrealistic to expect marginal-cost pricing. Consequently, a positive Lerner index would not necessarily reflect market power.

* Corresponding author.
E-mail addresses: l.spierdijk@rug.nl (L. Spierdijka), M.Zaouras@rug.nl (M. Zaourasa).

[1] There are also market imperfections under which marginal cost pricing is observed, without attaining the social optimum (e.g., externalities). We do not consider such market imperfections in the present study.

https://doi.org/10.1016/j.jbankfin.2017.09.022
0378-4266/

The theoretical part of this study proposes a new, scale-corrected price-cost margin for firms whose output level is in the economies of scale range. The new measure is defined as the relative difference between the observed price and the minimum-average cost level. In this way, it corrects the divergence of the price from marginal cost for the restriction of output (referred to as the 'scale correction'). We formalize the aforementioned concerns in the literature by formulating conditions under which this new measure acts as a sharper upper bound for the degree of market power than the Lerner index. We also identify conditions under which the scale-corrected Lerner index is required to get a proper upper bound on the degree of market power in the presence of economies of scale; a case not previously addressed in the literature.

The limitation of the Lerner index with respect to economies of scale seems particularly relevant for the U.S. banking sector. Recent literature has provided robust empirical evidence of economies of scale in this industry (Feng and Serletis, 2010; Wheelock and Wilson, 2012; Hughes and Mester, 2013; Beccalli et al., 2015). The empirical part of this study therefore considers U.S. commercial banks during the 2000 – 2014 period. Our analysis confirms the prior evidence on the ample presence of economies of scale for banks of all sizes. We find that, for 15–40% of the bank-years in our sample, the scale-corrected price-cost margin is smaller than the conventional Lerner index. This difference is (both statistically and economically speaking) significant, showing that these banks have less market power than suggested by the Lerner index. For the remaining 60–85% of the bank-years, the scale-adjusted Lerner index is significantly higher than the Lerner index, providing a proper upper bound on the degree of market power, unlike the Lerner index itself. Throughout, the magnitude of the scale-corrected Lerner index is substantial, suggesting significant market power among U.S. commercial banks. This is important information for regulators and policymakers, because it refutes the conjecture that deviations from marginal-cost pricing in the U.S. banking sector are merely due to scale economies.

The scale correction proposed in this study bears some resemblance to the efficiency correction proposed by Koetter et al. (2012) and Kutlu and Sickles (2012). However, while they focus on improving the Lerner index as a measure of the social loss (by adding a component to the conventional Lerner index related to ignored inefficiencies), we focus on improving the Lerner index as a measure of market power.

The setup of the remainder of this study is as follows. We introduce our theoretical framework and the scale-corrected Lerner index in Section 2, where we also analyze its properties. We provide an empirical illustration involving U.S. commercial banks in Section 3. Finally, Section 4 concludes.

2. A scale-corrected price-cost margin

This section introduces a new scale-corrected price-cost margin. Although our analysis is fully general (and thereby applicable to firms of any sort), we will formulate it in terms of banks for the sake of exposition. We distinguish two different cases on the basis of banks' cost function and output level, which we discuss in separate subsections.

2.1. Case 1: increasing marginal costs

We start with the case of increasing MC.

2.1.1. Assumptions and definitions

We consider banks with a single-output production technology. The associated cost function is written as $C(q)$, where q denotes a bank's output level. The marginal cost (MC) and average cost (AC) functions are denoted $MC(q) = C'(q)$ and $AC(q) = C(q)/q$, respectively.

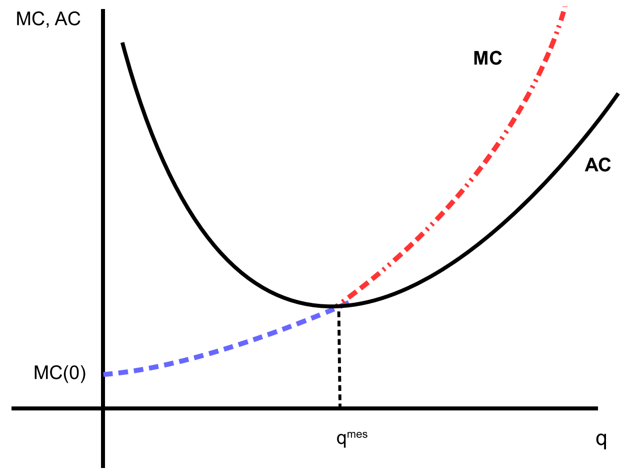

Fig. 1. Price feasibility and competitiveness.
Notes: Competitive prices are at or above $MC(0)$ and feasible prices at or above $AC^{\min}$. Any price at or above $AC^{\min}$ is both feasible and competitive.

A bank's Lerner index is defined as the bank's relative markup of the market output price (P^*) over marginal cost ($MC(q)$), given the bank's output level $q > 0$:[2]

$$L(q) = \frac{P^* - MC(q)}{P^*}. \quad (1)$$

We will refer to $MC(q)$ as the Lerner index' 'benchmark' price. We will need a few definitions to analyze the properties of a given benchmark price:

Definition 2.1.

(i) A price P is called *feasible* if $P \geq AC(q)$ for some $q > 0$.
(ii) A price P is called *competitive* if $P = MC(q)$ for some $q > 0$.
(iii) A price P is called *feasible and competitive* if $P = MC(q) \geq AC(q)$ for some $q > 0$.

Throughout, we will make one or more of the following assumptions:

Assumption 2.1.

(i) The AC-function is intersected by the MC cost function exactly once, in $q^{MES} = \arg\min_{q>0} AC(q)$. The output level q^{MES} is referred to as the minimum-efficient scale (MES), with $AC'(q^{MES}) = 0$, $AC'(q) < 0$ for $q < q^{MES}$ and $AC'(q) > 0$ for $q > q^{MES}$.
(ii) The observed output level q falls in the economies of scale range ($q < q^{MES}$);
(iii) Marginal-cost pricing is infeasible at the observed output level ($MC(q) < AC^{\min}$ for $q < q^{MES}$).

Fig. 1 illustrates the concepts of price feasibility and competitiveness under Assumption 2.1 (i) in case of increasing MC and U-shaped AC.

2.1.2. The Lerner index

A positive Lerner index is commonly associated with market power (Lerner, 1934). Yet a positive the Lerner index does not necessarily imply the presence of market power. This is particularly

[2] In the presence of multiple banks, the inverse market demand function evaluated in the total market output (denoted $P(Q^*)$) should be used in the Lerner index. For simplicity of notation, however, we use the shorthand notation P^* for the observed market price.

evident under Assumption 2.1. Because $MC(q) < AC^{\min} < AC(q)$ under Assumption 2.1(iii), marginal-cost pricing would result in negative profits for the bank. Hence, if we observe a positive Lerner index in such a case, it is intuitively clear that part of this positive Lerner index will merely be due to the infeasibility of the benchmark price instead rather than to market power.[3] The Lerner index therefore acts as a strict upper bound on the degree of market power.

2.1.3. A scale-corrected price-cost margin: intuition

Section 2.1.2 has made clear that, under Assumption 2.1, part of the positive Lerner index merely reflects the infeasibility of MC-pricing due to scale economies. We are therefore interested in finding an alternative benchmark price in order to 'repair' the Lerner index in such cases. We will do so by proposing an upward correction of the benchmark price to account for the infeasibility of MC-pricing. The result is a sharper upper bound on the degree of market power than given by the original Lerner index.

So the main question is: what is an appropriate alternative benchmark price? Because of its feasibility, $AC(q)$ may seem a good alternative benchmark price. However, we have already known since (Lerner, 1934) that the profit margin is not a good measure of market power. We therefore have to find another alternative benchmark price.

A necessary condition for the benchmark price is feasibility and competitiveness in the sense of Definition 2.1(iii). It should be feasible to make sure that the benchmark price is actually attainable by banks in practice, and competitive because we aim at measuring market power. However, there are many prices that are feasible and competitive, as we can see in Fig. 1. So we are left with the question which one to use.

As can be seen from Fig. 1, $AC^{\min} = AC(q^{MES}) = MC(q^{MES})$ is the *lowest* feasible and competitive price. $AC^{\min}$ arises in the hypothetical competitive equilibrium where many price-taking banks produce quantity q^{MES} and make zero profit by offering their good against price $AC^{\min}$. The alternative price-cost margin that we propose is obtained by replacing the benchmark price in the Lerner index by $AC^{\min}$. The resulting scale-corrected price-cost margin is thus defined as

$$L^{AC^{min}} = \frac{P^* - AC^{min}}{P^*}. \quad (2)$$

and reflects the *maximum* relative difference between the observed price and a feasible and competitive price. The properties of $L^{AC^{\min}}$ will be explored in the next section.

2.1.4. Properties of the scale-corrected price-cost margin

Under Assumption 2.1, we can make the following decomposition of the numerator of $L^{AC^{\min}}$:

$$P^* - AC^{min} = \underbrace{P^* - MC(q)}_{(1)} - \underbrace{\left[AC^{min} - MC(q)\right]}_{(2)}. \quad (3)$$

The decomposition is visualized in Fig. 2 for the usual case that P^* is a feasible price. Component (1) is the numerator of the Lerner index. Component (2) is positive and subtracted from the Lerner index to correct for the fact that a price below $AC^{\min}$ can *never* be feasible and competitive. Hence, part of the Lerner index is unlikely to be due to market power, and $L^{AC^{\min}} \leq L$ corrects for this. We thus see that, unlike the Lerner index, $L^{AC^{\min}}$ explicitly accounts

[3] Instead of saying that 'marginal-cost pricing is infeasible in the presence of economies of scale', we could also opt for a more formal formulation: with economies of scale, competitive equilibrium (consisting of marginal-cost pricing and a zero-profit requirement) is impossible without lump-sum transfers (Baumol, 1979).

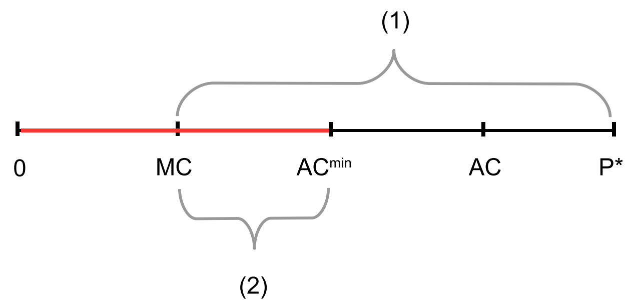

Fig. 2. Decomposition of $P^* - MC(q)$.
Notes: The red horizontal line indicates the range where prices are infeasible. Component (1) is the deviation of the price from MC; Component (2) is the scale correction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

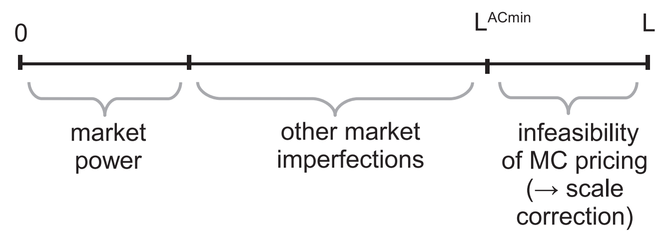

Fig. 3. L and $L^{AC^{min}}$ and market power, other market imperfections and scale correction.
Notes: This decomposition divides the social loss reflected by the Lerner index into three parts: market power, infeasibility of MC-pricing, and remaining market imperfections.

for the restriction of output via Component (2) (the 'scale correction'). The degree of output restriction is reflected by the difference between the observed output level q and the minimum-efficient scale q^{MES}. Stated differently, $L^{AC^{\min}}$ is the relative price difference between two *different* output levels (the bank's observed output quantity and the minimum-efficient scale), while the Lerner index is the relative difference between two possible prices for the *same* output level (the bank's observed output quantity). Hence, $L^{AC^{\min}}$ is a *two-price-two-output-level* measure, while the Lerner index is a *two-price-one-output-level* measure. $L^{AC^{\min}}$ accounts for both the divergence of the price from marginal costs and the restriction of output.

Fig. 3 provides a stylized decomposition of the Lerner index and our new index in terms of the social loss due to market power, infeasibility of MC-pricing (scale correction), and remaining market imperfections under Assumption 2.1. We observe that the Lerner index captures the social loss due to deviations from MC-pricing and acts as an upper bound on the degree of market power. By correcting for scale, $L^{AC^{\min}}$ provides a sharper upper bound on the degree of market power than the Lerner index itself.

2.2. Case 2: U-shaped marginal costs

We now turn to the case of U-shaped MC with low output levels.

2.2.1. Extension to U-shaped MC and diseconomies of scale

In case of U-shaped MC, Assumption 2.1 (iii) will not hold for sufficiently small output levels $0 < q < q^{MES}$. More formally, assume that Assumptions 2.1 (i) and (ii) hold, but that $MC(q) > AC^{\min}$ for a certain bank with a relatively small output level $q < q^{MES}$. As shown in Fig. 4, $MC(q) = MC(q^*)$. This means that, according to Definition 2.1 (iii), $MC(q)$ is actually a *feasible and competitive* price. As before, we expect the Lerner index to be non-negative, since banks will avoid making losses. A positive distance to a feasible

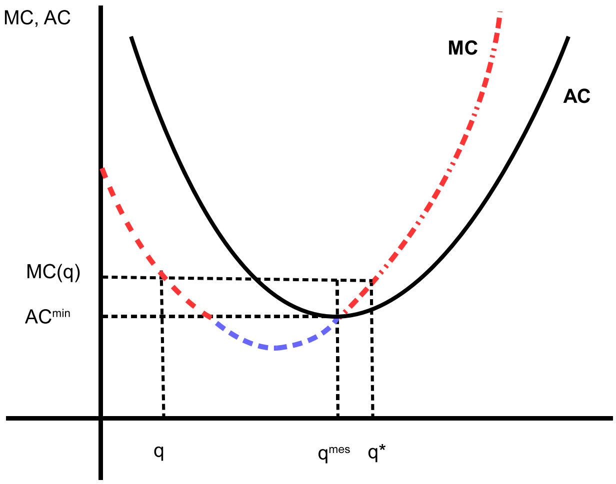

Fig. 4. Price feasibility and competitiveness (U-shaped MC).
Notes: Any price at or above $AC^{\min}$ is both feasible and competitive.

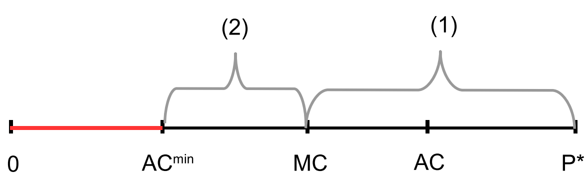

Fig. 5. Decomposition of $P^* - AC^{\min}$.
Notes: The red horizontal line indicates the range where prices are infeasible. Component (1) is the deviation of the price from MC; Component (2) is the profit correction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and competitive price suggests the presence of market power.[4] But how much market power?

Because positive profits can be earned with price $MC(q)$, the latter price is relatively high. Therefore, the gap between an observed price P^* and the benchmark price $MC(q)$ is relatively small. Consequently, a low value of the Lerner index may be due to the high profits that can be earned with $MC(q)$ rather than to the competitiveness of the price P.[5] We therefore have to make a profit correction to the Lerner index to remove this effect. Again we do this by looking at $L^{AC^{\min}}$ instead of L. To see how the latter measure applies a profit correction to the Lerner index, we consider the following decomposition:

$$P^* - AC^{min} = \underbrace{P^* - MC(q)}_{(1)} + \underbrace{\left[MC(q) - AC^{min}\right]}_{(2)}. \quad (4)$$

The decomposition is visualized in Fig. 5 for the usual case that P^* is a feasible price. Component (1) is the numerator of the Lerner index. Component (2) is positive and added to the numerator of the Lerner index to correct for the fact that the benchmark price $MC(q)$ has an "excess profit" of $MC(q) - AC^{\min}$ relative to the lowest feasible and competitive (and unique zero-profit) price $AC^{\min}$. Although Component (2) acts as a profit correction, it is also a scale correction: it arises because $q < q^{MES}$. We therefore continue to refer to $L^{AC^{\min}}$ as the "scale-corrected" Lerner index.

Fig. 6 provides a stylized decomposition of the Lerner index and our new index in terms of the social loss due to market power, remaining market imperfections and the profit correction. $L^{AC^{\min}}$ provides a proper upper bound on the degree of market power, unlike the Lerner index itself.

Using exactly the same arguments, we must consider $L^{AC^{\min}}$ for banks facing diseconomies of scale.

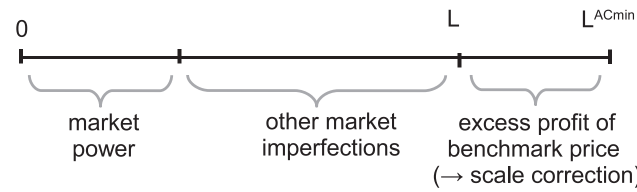

Fig. 6. L and $L^{AC^{\min}}$ and market power, other market imperfections and profit correction.

3. Empirical illustration

This section investigates the market power of U.S. commercial banks during the period 2000 – 2014. The U.S. banking sector is particularly relevant in our setting, because of the recent attention for economies of scale in banking and the robust evidence for it (e.g. Feng and Serletis, 2010; Wheelock and Wilson, 2012; Hughes and Mester, 2013; Beccalli et al., 2015; Feng and Zhang, 2014).

3.1. Translog cost function

We will use the popular translog cost function to capture U.S. commercial banks' total costs as a function of output (see e.g. Koetter et al., 2012). As a first step, we briefly analyze the properties of this cost function in terms of the shape of the AC and MC-functions. Without loss of generality, we confine the analysis to the generic translog cost function given by

$$\log(C(q)) = a + b\log(q) + (c/2)\log(q)^2, \quad (5)$$

where $q > 0$ denotes the bank's output level. The AC-function then equals

$$AC(q) = \exp\big(a + (b-1)\log(q) + (c/2)\log(q)^2\big). \quad (6)$$

The AC-function is U-shaped for $c > 0$. In this case we find q^{MES} by solving $\partial AC(q)/\partial q = 0$, or equivalently, $\partial\log(C(q))/\partial\log(q) = 1$. This yields $q^{MES} = \exp((1-b)/c)$ and $AC^{\min} = AC(q^{MES})$. The MC-function equals $MC(q) = [b + c\log(q)]AC(q)$, showing that $MC(q) \to -\infty$ for $q \to 0$. The first derivative of the MC-function equals

$$MC'(q) = (AC(q)/q)\underbrace{[(b-1)b + c + c\log(q)(2b - 1 + c\log(q))]}_{(*)} \quad (7)$$

The part of $MC'(q)$ that is marked with a (*) is a downward-sloping parabola in $\log(q)$ for $q > 0$. Hence, the MC-function is either increasing or nonmonotonic.[6] In both cases the MC-function intersects the AC curve in q^{MES}.

[4] There is also a more formal reason to expect market power in this case. Assuming that banks maximize profits, the observed output level q in Fig. 4 must correspond to a downward-sloping demand curve. That is, with a flat demand curve p the profit-maximization problem $p = MC$ has two possible solutions due to the U-shaped MC curve: q such that $p = MC(q)$ and $q^* > q^{MES} > q$ such that $p = MC(q^*)$. Since the former solution results in negative profit and the latter solution in positive profit, the profit-maximizing output level is q^*. Assuming that banks maximize profits, we conclude that cannot have a flat demand curve if their observed output level is q. This means that banks must have some market power.

[5] More formally, a benchmark price yielding excess profit cannot correspond to a sustainable competitive outcome; that is, competitive equilibrium (defined as marginal-cost pricing and a zero-profit requirement) is impossible without taxing away excess profit (Baumol, 1979).

[6] The MC-function is increasing for $\tilde{a} > \tilde{b}^2/(4\tilde{c})$, with $\tilde{a} = ((b-1)b + c)(2b-1)$, $\tilde{b} = c(b^2 + b + c - 1)$, $\tilde{c} = c^2$. It is nonmonotonic for $\tilde{a} < \tilde{b}^2/(4\tilde{c})$. In case of nonmonotonicity, the shape of the MC-function is increasing-decreasing-increasing. For values of q sufficiently far away from 0, the MC curve is then U-shaped.

Table 1
Sample statistics.

banks that are part of a BHC				
	2000 – 2007		2008 – 2014	
	mean	*s.e.*	*mean*	*s.e.*
price of purchased funds	4.1%	1.4%	2.1%	1.2%
price of core deposits	2.2%	1.2%	0.9%	0.7%
wage rate	48.7	15.0	62.8	18.6
price of physical capital	33.1%	30.5%	34.4%	40.4%
total assets	1,279,105	21,041,907	2,416,749	43,160,688
total costs	50,769	843,277	56,226	955,480
total costs/total assets	4.3%	1.1%	3.0%	1.1%
operating income/total assets	6.9%	2.1%	5.3%	1.5%
# bank-years	47,713		33,992	
# years	8		7	
# banks	7514		5874	
stand-alone banks				
	2000 – 2007		**2008 – 2014**	
	mean	*s.e.*	*mean*	*s.e.*
price of purchased funds	3.7%	1.5%	1.9%	1.1%
price of core deposits	2.2%	1.1%	1.0%	0.8%
wage rate	51.1	17.3	69.2	23.2
price of physical capital	42.4%	40.5%	51.5%	59.8%
total assets	141,696	744,407	229,076	1,213,212
total costs	5784	26,745	6060	20,791
total costs/total assets	4.4%	1.4%	3.1%	1.4%
operating income/total assets	6.6%	3.1%	5.0%	2.2%
# bank-years	10,315		6374	
# years	8		7	
# banks	2165		1217	

Notes: The columns captioned 'mean' report the sample means, while the columns captioned 's.e.' show the sample standard error. All level variables are in thousands of $.

3.2. Data

We assume that banks employ a production technology with four inputs and one output factor (cf. Wheelock and Wilson, 2012). The choice for inputs and outputs is based on a modernized version of the intermediation model for banking (Klein, 1971; Monti, 1972). The four inputs we consider are purchased funds, core deposits, labor services, and physical capital. The corresponding input prices are (1) the price of purchased funds of bank $i = 1, \ldots, N$ in year $t = 1, \ldots, T$ ($P_{1,\,it}$), (2) the core deposit interest rate ($P_{2,\,it}$), (3) the wage rate ($P_{3,\,it}$), and (4) the price of physical capital ($P_{4,\,it}$). Total operating costs (C_{it}) are defined as the sum of expenses on purchased funds, core deposits, personnel expenses, and expenses on physical capital. The single output factor we consider is total assets (Y_{it}).[7] We follow Koetter et al. (2012) and proxy the output price by average revenue, which is the ratio of operating income to total assets. The analysis below is easily corrected to the case of less, more or different input factors.

We use year-end Call Report Data to create an annual sample of U.S. commercial banks covering the years 2000 – 2014 containing the above variables. Appendix A explains how the Call Report Data have been used to obtain the required variables. We restrict our sample to commercial banks with a physical location in a U.S. state, subject to deposit-related insurance. We split up the sample in two subperiods (2000 – 2007 and 2008 – 2014) to account for structural change around the start of the crisis. We apply trimming to both samples to get rid of outliers and we also remove bank-year observations with inconsistent values.

Table 1 provides separate sample statistics for the pre-crisis and (post-) crisis periods. Furthermore, we also report separate sample statistics for banks that are part of a bank-holding company (BHC) and stand-alone banks. These two groups of banks will come back later in our empirical analysis and we therefore consider each group separately. This results in four subsamples: pre-crisis/banks that are part of a BHC, pre-crisis/stand-alone banks, (post-) crisis/banks that are part of a BHC, and (post-) crisis/stand-alone banks.

Throughout, we observe a substantial decline in the prices of purchased funds and core deposits after the onset of the crisis, which reflects the actions taken by the Fed to boost the U.S. economy. Furthermore, the (post-) crisis sample is characterized by a larger average bank size due to the consolidation wave that took place after the onset of the crisis (Dunn et al., 2015). We additionally observe a higher average wage rate and lower mean average costs and revenues after the onset of the crisis. Moreover, there are about five times as much bank-years corresponding to BHCs than bank-years associated with stand-alone banks. In terms of average total assets, banks that are part of a BHC are about ten times larger than stand-alone banks.

3.3. Model specification

In this empirical example we use a more complex translog cost function than in Section 3.1. As usual, we impose linear homogeneity in input prices by normalizing total costs and input prices with the price of purchased funds ($P_{1,\,it}$). Throughout, variables with a tilde have been normalized with the price of purchased funds prior to taking the logarithmic transformation to ensure linear homogeneity. This results in the following total cost regression model for bank i in year t:

$$\log(\widetilde{C}_{it}) = \alpha_i + \sum_{j=2}^{4} \beta_{j,p} \log(\widetilde{P}_{j,it}) + (1/2) \times \sum_{j=2}^{4}\sum_{k=2}^{4} \beta_{jk,pp} \log(\widetilde{P}_{j,it}) \log(\widetilde{P}_{k,it})$$

[7] The notation in this section is in line with empirical banking studies, but differs from the notation used in our theoretical analysis.

$$+\sum_{j=2}^{4}\beta_{j,py}\log(\widetilde{P}_{j,it})\log(Y_{it})+\beta_{y}\log(Y_{it})$$
$$+(1/2)\beta_{yy}\log(Y_{it})^2$$
$$+\beta_{time,y}t\log(Y_{it})+\beta_{time2,y}t^2\log(Y_{it})+\beta_{time}t$$
$$+\beta_{time2}t^2+\varepsilon_{it}, \quad (8)$$

where α_i is a bank-specific effect, t a time trend accounting for technological change, and ε_{it} a zero-mean error term that is orthogonal to the regressors. We notice that the functional form of the translog cost function in Eq. (8) is similar to the stylized one considered in Section 3.1, but the parameters a and b of Eq. (5) now depend on bank-specific and time-specific variables.[8] The MC-function associated to Eq. (8) equals

$$MC_{it}=\frac{C_{it}}{Y_{it}}\frac{\partial \log C_{it}}{\partial \log Y_{it}}$$
$$=\frac{C_{it}}{Y_{it}}\Big(\beta_y+\sum_{j=2}^{4}\beta_{j,py}\log(\widetilde{P}_{j,it})+\beta_{yy}\log(Y_{it})$$
$$+\beta_{time,y}t+\beta_{time2,y}t^2\Big). \quad (9)$$

The corresponding log AC-function is given by

$$\log(AC_{it})=\alpha_i+\log(\widetilde{P}_{1,it})+\sum_{j=2}^{4}\beta_{j,p}\log(\widetilde{P}_{j,it}) \quad (10)$$
$$+(1/2)\sum_{j=2}^{4}\sum_{k=2}^{4}\beta_{jk,pp}\log(\widetilde{P}_{j,it})\log(\widetilde{P}_{k,it})+\log(P_{1,it})$$
$$+\sum_{j=2}^{4}\beta_{j,py}\log(\widetilde{P}_{j,it})\log(Y_{it})+(\beta_y-1)\log(Y_{it})$$
$$+(1/2)\beta_{yy}\log(Y_{it})^2$$
$$+\beta_{time,y}t\log(Y_{it})+\beta_{time2,y}t^2\log(Y_{it})$$
$$+\beta_{time}t+\beta_{time2}t^2. \quad (10)$$

AC are U-shaped for $\beta_{yy}>0$, in which case the minimum-efficient scale is found by solving $\partial\log(C_{it})/\partial\log(Y_{it})=1$. This results in

$$Y_{it}^{MES}=\exp\Big(-\Big([\beta_y-1]+\sum_{j=2}^{4}\beta_{j,py}\log(\widetilde{P}_{j,it})+\beta_{time,y}t$$
$$+\beta_{time2,y}t^2\Big)/\beta_{yy}\Big). \quad (11)$$

We notice that Y_{it}^{MES} is unrestricted, in the sense that it does not necessarily lie strictly interior to the size range of the banks in the sample (Shaffer, 1998). Substituting Y_{it}^{MES} in the AC-function yields $AC_{it}^{\min}=AC_{it}(Y_{it}^{MES})$. For bank i in year t we calculate

$$L_{it}=(P_{it}-MC_{it})/P_{it},\qquad L_{it}^{AC^{\min}}=(P_{it}-AC_{it}^{min})/P_{it}. \quad (12)$$

It is important to notice that $AC^{\min}$, Y^{MES}, L and $L^{AC^{\min}}$ are both bank-specific and time-specific.

3.4. Estimation results

The literature has emphasized the relevance of bank-specific cost technologies (McAllister and McManus, 1993; Berger and Humphrey, 1997; Kumbhakar and Tsionas, 2008). Furthermore, as acknowledged by Feng and Serletis (2010) and Wheelock and Wilson (2012), the translog cost function of Eq. (8) is only suitable for samples consisting of relatively homogeneous banks. We deal with these issues in several ways. First, we follow Wheelock and Wilson (2012) by estimating separate translog cost functions for banks that are part of a BHC and stand-alone banks. Second, we do this for both the pre-crisis and (post-) crisis periods, resulting in four subsamples with separate estimation results. Third, in all four cases we estimate Eq. (8) using the random-effect estimator to capture any additional bank-specific heterogeneity.[9]

We confine the discussion of the estimated model coefficients to the observation that the estimates of β_{yy} are significantly positive at the 1% level and jointly significant at the 1% level with the estimate of β_y, based on clustered standard errors that account for time-series correlation and heteroskedasticity. During the pre-crisis period, the values of $\hat{\beta}_{yy}$ are 0.014 (banks that are part of a BHC) and 0.053 (stand-alone banks). During the (post-) crisis period these values are 0.019 and 0.040, respectively. This yields U-shaped AC in all four cases.

Tables 2 and 3 report sample statistics for L and $L^{AC^{\min}}$ in three ways: (1) for all bank-years with $MC<AC$, (2) for all bank-years with $MC<AC$ and $L^{AC^{\min}}<L$ (the case discussed in Section 2.1) and (3) for all bank-years with $MC<AC$ and $L^{AC^{\min}}>L$ (the case discussed in Section 2.2). The percentage of bank-years that falls in each of the three categories is also reported (see the rows with the caption '% of bank-years'). Throughout, Tables 2 and 3 report sample medians and interquartile ranges (IQR) instead of sample means and standard errors to avoid that a few outliers contaminate the results.[10]

Tables 2 and 3 report economies of scale for more than 95% of all bank-years in each subsample, confirming the relevance of our analysis. For banks that are part of a BHC, about 20% (pre-crisis) and 15% ((post-) crisis) of the bank-year observations with economies of scale result in $L^{AC^{\min}}<L$. For the remaining 80% and 85% of the bank-years, we find $L^{AC^{\min}}>L$. For stand-alone banks, about 40% (pre-crisis) and 20% ((post-) crisis) of the bank-year observations with economies of scale result in $L^{AC^{\min}}<L$. For the remaining 60% and 80% of the bank-years, we find $L^{AC^{\min}}>L$.

Our interpretation of the results focuses on the difference between the estimated values of L and $L^{AC^{\min}}$. As a first step, we conduct a Wilcoxon rank test to test whether the difference between L and $L^{AC^{\min}}$ is statistically speaking significant. This test can be viewed as a "non-parametric paired t-test" and allows us to assess whether the location parameter of the distribution of $L-L^{AC^{\min}}$ is equal to $\mu=0$. The null hypothesis is $H_0:\mu=0$, while the alternative hypothesis is $H_1:\mu\neq 0$. For each of the four subsamples, we reject H_0 with p-value 0. The test outcomes show that, in terms of the median, the paired difference between L and $L^{AC^{\min}}$ is significantly different from 0. Now that we have established the statistical significance of the difference between L and $L^{AC^{\min}}$, we can analyze the economic relevance of this difference.

By comparing the medians of L and $L^{AC^{\min}}$ in Tables 2 and 3, we conclude that the magnitude of this difference ranges between 5.2 and 12.7 percentage points. To assess the economic relevance of this difference, Tables 2 and 3 also consider the relative difference $(L-L^{AC^{\min}})/L$. For the bank-year observations with $MC<AC$ and $L^{AC^{\min}}<L$, this relative difference reflects the percentage of the Lerner index that is due to the infeasibility of marginal-cost pric-

[8] More precisely, we now have $a=\alpha_i+\sum_{j=2}^{4}\beta_{j,p}\log(\widetilde{P}_{j,it})+(1/2)\sum_{j=2}^{4}\sum_{k=2}^{4}\beta_{jk,pp}\log(\widetilde{P}_{j,it})\log(\widetilde{P}_{k,it})+\beta_{time}t+\beta_{time2}t^2$ and $b=\beta_y+\sum_{j=2}^{4}\beta_{j,py}\log(\widetilde{P}_{j,it})+\beta_{time,y}t+\beta_{time2,y}t^2$.

[9] We do not use the FE estimator because total assets has relatively little time variation, due to which part of the effect of (log) total assets on (log) total costs might be absorbed by the fixed effect.

[10] Such outliers may arise when the denominator of the (scale-corrected) Lerner index is relatively small.

Table 2
Summary of estimation results (banks that are part of a BHC).

	2000 – 2007		2008 – 2014	
	bank-years with $MC < AC$			
% bank-years	98.5%		99.3%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	23,128,210	41,914,474	67,549,684	42,502,568
L	41.4%	11.2%	49.6%	11.8%
$L^{AC^{\min}}$	48.4%	14.9%	59.9%	13.1%
$(L - L^{AC^{\min}})/L$	-16.3%	28.3%	-21.3%	28.4%
5% quantile of $(L - L^{AC^{\min}})/L$	-57.5%		-65.5%	
95% quantile of $(L - L^{AC^{\min}})/L$	19.8%		14.0%	
	bank-years with $MC < AC$ **and** $L^{AC^{\min}} < L$			
% bank-years	21.1%		14.9%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	25,433,776	42,930,185	65,443,240	39,985,875
L	43.2%	12.5%	55.4%	12.2%
$L^{AC^{\min}}$	38.0%	13.5%	49.2%	14.2%
$(L - L^{AC^{\min}})/L$	9.9%	14.7%	9.2%	13.7%
95% quantile of $(L - L^{AC^{\min}})/L$	41.6%		42.3%	
	bank-years with $MC < AC$ **and** $L^{AC^{\min}} > L$			
% bank-years	77.4%		84.4%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	14,918,685	33,721,003	81,205,227	53,714,843
L	41.0%	10.9%	48.7%	11.3%
$L^{AC^{\min}}$	50.8%	13.1%	61.4%	11.7%
$(L - L^{AC^{\min}})/L$	-22.0%	23.1%	-25.3%	24.6%
95% quantile of $(L - L^{AC^{\min}})/L$	-61.8%		-69.0%	

Notes: The columns captioned “median” report the sample median, while the columns captioned “IQR” show the sample interquartile range. Y^{MES} is expressed in thousands of $.

Table 3
Summary of estimation results (stand-alone banks).

	2000 – 2007		2008 – 2014	
	bank-years with $MC < AC$			
% bank-years	95.6%		98.2%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	878,037	1,044,292	8,663,497	10,568,698
L	43.1%	13.4%	50.8%	13.8%
$L^{AC^{\min}}$	44.9%	18.8%	59.6%	17.7%
$(L - L^{AC^{\min}})/L$	-5.9%	31.5%	-19.2%	30.0%
5% quantile of $(L - L^{AC^{\min}})/L$	-53.7%		-69.7%	
95% quantile of $(L - L^{AC^{\min}})/L$	36.2%		17.4%	
	bank-years with $MC < AC$ **and** $L^{AC^{\min}} < L$			
% bank-years	38.0%		18.1%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	930,205	1,196,418	8,700,024	10,738,663
L	43.7%	12.6%	54.5%	12.8%
$L^{AC^{\min}}$	36.3%	14.7%	48.1%	14.0%
$(L - L^{AC^{\min}})/L$	14.1%	19.0%	9.0%	14.5%
95% quantile of $(L - L^{AC^{\min}})/L$	50.8%		42.9%	
	bank-years with $MC < AC$ **and** $L^{AC^{\min}} > L$			
% bank-years	57.6%		80.1%	
	median	*IQR*	*median*	*IQR*
Y^{MES}	815,656	877,363	8,363,944	10,179,269
L	42.8%	14.0%	50.0%	13.8%
$L^{AC^{\min}}$	50.8%	16.2%	62.0%	15.9%
$(L - L^{AC^{\min}})/L$	-18.1%	22.5%	-24.2%	26.2%
95% quantile of $(L - L^{AC^{\min}})/L$	-68.1%		-76.2%	

Notes: The columns captioned “median” report the sample median, while the columns captioned “IQR” show the sample interquartile range. Y^{MES} is expressed in thousands of $.

ing. For the bank-year observations with $MC < AC$ and $L^{AC^{\min}} > L$, the relative difference reflects the percentage by which the Lerner index should be increased to get a proper upper bound on the degree of market power. Tables 2 and 3 show that median relative difference is economically substantial and can amount to as much as 14.1% ($L^{AC^{\min}} < L$; stand-alone banks; pre-crisis) and -25.3% ($L^{AC^{\min}} > L$; banks that are part of a BHC; post-crisis). For certain bank-years the relative difference is even higher, as shown by the sample quantiles of $(L - L^{AC^{\min}})/L$ as reported Tables 2 and

3. In sum, the difference between L and $L^{AC^{\min}}$ is not only statistically, but also economically significant.[11]

Despite the difference in cost technology, we observe little difference in the median values between the (scale-corrected) Lerner indices of banks that are part of a BHC and stand-alone banks. Furthermore, we see that the medians of the (scale-corrected) Lerner index are higher during the (post-) crisis period than during the pre-crisis period. Given that we use an unbalanced sample of banks, one reason to expect higher market power during the (post-) crisis period is the possibility that only banks' with relatively high market power managed to survive the crisis. We also notice that our values of the Lerner index are somewhat higher than documented in other U.S. Lerner studies, such as Koetter et al. (2012), who consider the period 1976 – 2007.

3.5. Extension to the multi-output case

Our theoretical and empirical analyses have assumed a single-output cost technology. Extension to a multi-output setting with $K \geq 2$ output factors is theoretically straightforward, but empirically difficult. The calculation of output-specific minimum average costs would require knowledge the AC of K different output factors. For these AC to be defined, a bank's cost technology has to be additively separable, meaning that a bank's total costs C should be decomposable as $C = C_1 + \ldots + C_K$ (Demski, 2008, pp. 44–45). The AC of output k would then equal C_k/Y_k, with Y_k the total value of output $k = 1, \ldots, K$. It is unrealistic to assume additive separability, however, since some input factor costs (like labor and capital) will apply to multiple outputs. Furthermore, even under the assumption of additive separability, we would still face the challenging task of determining which input factor costs correspond to which output. Our data do not provide this information and no empirical strategy exists for matching input factor costs with output factors. Our analysis is therefore based on a single-output cost function that has been widely applied in the banking literature; for some recent examples of similar cost function we refer to e.g. Koetter et al. (2012), Kasman and Carvallo (2014), Mirzaei and Moore (2014) and Fu et al. (2014).

This raises the question whether the focus on a single-output cost technology is a serious limitation of our analysis. Because output-specific marginal costs for any K-output cost technology are defined as the partial derivative of the total cost function, the conventional Lerner index is easily obtained for each output (provided that we also have output-specific prices or proxies of those prices). This allows us to assess whether our single-output Lerner index is some sort of weighted average of the output-specific Lerner indices, in which case our analysis would still have a decent economic interpretation.

We have verified this for our sample of banks. We choose total loans and total securities as the two output factors (with the same input factors as before) and we proxy the two output prices by “interest and fee income on loans” divided by total loans and “operating income minus interest and fee income on loans” divided by total securities, respectively. This results in two Lerner indices: one for total loans and one for total securities. The medians of the two Lerner indices in the 2-output model are about 0.40 (pre-crisis period) and 0.50 ((post-) crisis period) for total loans and 0.50 (pre-crisis period) and 0.60 ((post-) crisis period) for total securities. We thus see that the median values of the single-output Lerner index are close to the median values of the 2-output Lerner index corresponding to total loans. The explanation is that total loans have a much larger average output share (about 75%) than total securities (about 25%) and are thus the main determinant of the single-output Lerner index. This suggests that our single-output analysis, comprising both conventional and scale-corrected Lerner indices, still has a sound economic interpretation: it is representative for the loan-part of total assets.

[11] We notice that we obtain qualitatively similar results if we estimate a single translog cost function for banks that are part of a BHC and stand-alone banks instead of two separate translog cost functions.

4. Conclusion

The theoretical part of this study has proposed a new, scale-corrected price-cost margin for firms that produce in the economies of scale range. The new measure is defined as the relative difference between the observed price and the minimum-average cost level. In this way, it corrects the divergence of the price from marginal costs for the restriction of output (referred to as the ‘scale correction’). We have formulated conditions under which this new measure acts as a sharper upper bound for the degree of market power than the Lerner index. We have also formulated conditions under which the scale-corrected Lerner index is required to get a proper upper bound on the degree of market power.

The empirical part of this study has analyzed the degree of market power in the U.S. banking sector during the 2000 – 2014 period. This sector is particularly relevant, because the recent literature has provided robust empirical evidence of economies of scale in this sector. Our analysis has confirmed this evidence and has shown that, for 15–40% of the bank-years in our sample, the scale-corrected price-cost margin is smaller than the Lerner index. This difference is both economically and statistically speaking significant, showing that these banks have less market power than suggested by the conventional Lerner index. For the remaining 60–85% of the bank-years, the scale-adjusted Lerner index is significantly higher than the Lerner index, providing a proper upper bound on the degree of market power, unlike the Lerner index itself. Throughout, the magnitude of the scale-corrected Lerner index is substantial, suggesting significant market power among U.S. commercial banks. In the light of the frequently expressed concerns in the economic literature that a positive Lerner index would not necessarily reflect market power in the presence of economies of scale, this is important information for regulators and policy-makers.

Our results apply to one specific sample of banks. We may find different results for other industries or time periods. We leave this as a topic for future research.

Acknowledgments

The authors are grateful to the participants of the 2016 Banking Conference at the Université d'Auvergne in Clermont-Ferrand, the seminar attendants at the University of Cyprus, and two anonymous referees and an associate editor for useful comments. Laura Spierdijk gratefully acknowledges financial support by a Vidi grant (452.11.007) in the 'Vernieuwingsimpuls' program of the Netherlands Organization for Scientific Research (NWO). The usual disclaimer applies.

Appendix A. Call report data

Table A.1 explains how the Call Report Data have been used to define the variables used in the empirical part of this study.

Table A1
Definition of variables.

Variable	Series
total loans	RCFD1400
securities	RCFD1754+RCFD1773
purchased funds	RCON2604+RCFD2800+RCFD3548+RCFD3200+RCFD3190+RCFD2200-RCON2200
core deposits	RCON2200-RCON2604
# of full-time equivalent employees	RIAD4150
physical capital	RCFD2145
expenditures on purchased funds (interest)	RIAD4172+RIAD4180+RIADA517+ RIAD4185+RIAD4200
expenditures on core deposits (interest)	RIAD4170-RIADA517- RIAD4172
expenditures on labor services (salaries)	RIAD4135
expenditures on physical capital	RIAD4217
total costs	sum of all expenditures
price of purchased funds	(expenditures on purchased funds)/(purchased funds)
core deposit rate	(expenditures on core deposits)/(core deposits)
wage rate	(expenditures on labor services)/(# of full-time equivalent employees)
price of physical capital	(expenditures on physical capital)/(physical capital)
total assets	RCFD2170
total interest and fee income on loans	RIAD4010
total interest income	RIAD4107
total non-interest income	RIAD4079
total operating income	total interest income + total non-interest income
is bank part of BHC?	RSSD9364
bank index	RSSD9050
time index	RSSD9999

Notes: This table explains how the variables in this study have been calculated from the data available in the Call Reports.

References

Amoroso, L., 1933. La curva statica di offerta. Giornale Degli Economisti 45, 126.
Amoroso, L., 1938. Principii di economica corporativa. Zanichelli, Bologna.
Amoroso, L., 1954. The static supply curve. In: Peacock, A., Stolper, W., Turvey, R., Henderson, E. (Eds.), International Economic Papers (English translation of Amoroso 1930), 4. McMillan, London, pp. 39–65.
Baumol, W.J., 1979. Quasi optimality: the price we must pay for a price system. J. Polit. Econ. 87, 578–599.
Beccalli, E., Anolli, M., Borello, G., 2015. Are European banks too big? Evidence on economies of scale. J. Bank Financ. 58, 232–246.
Berger, A., Humphrey, D., 1997. Efficiency of financial institutions: international survey and directions for future research. Eur. J. Oper. Res. 98, 175–212.
Blair, R., Sokol, D., 2014. The Oxford Handbook of International Antitrust Economics, 1. Oxford University Press, USA.
Cairns, R., 1995. Reflections on Lerner's index of monopoly power. Rev. Indus. Organ. 10, 83–96.
Demski, J., 2008. Managerial uses of accounting information, 2nd edition New York, NY: Worth Publishers.
Dunn, J., Intintoli, V., McNutt, J., 2015. An examination of non-government-assisted U.S. commercial bank mergers during the financial crisis. J. Econ. Bus. 77, 16–41.
Elzinga, K., Mills, D., 2011. The Lerner index of monopoly power: origins and uses. Am. Econ. Rev. 101, 558–564.
Feng, G., Serletis, A., 2010. Efficiency, technical change, and returns to scale in large U.S. banks: panel data evidence from an output distance function satisfying theoretical regularity. J. Bank. Financ. 34, 127–138.
Feng, G., Zhang, X., 2014. Returns to scale at large banks in the U.S.: a random coefficient stochastic frontier approach. J. Bank. Financ. 39, 135–145.
Fu, X., Lin, Y., Molyneux, P., 2014. Bank competition and financial stability in Asia pacific. J. Bank. Financ. 38, 64–77.
Fernández de Guevara, J., Maudos, J., 2004. Measuring welfare loss of market power: an application to European banks. Appl. Econ. Lett. 11, 833–836.
Giocoli, N., 2012. Who invented the Lerner Index? Luigi Amoroso, the dominant firm model, and the measurement of market power. Rev. Indus. Organ. 41.
Hughes, J., Mester, L., 2013. Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. J. Financ. Intermed. 22, 559–585.
Kasman, A., Carvallo, C., 2014. Financial stability, competition and efficiency in Latin American and Caribbean banking. J. Appl. Econ. 17, 301–324.
Klein, M., 1971. A theory of the banking firm. J. Money Credit Bank. 3, 205–218.
Koetter, M., Kolari, J., Spierdijk, L., 2012. Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for U.S. banks. Rev. Econ. Stat. 94, 462–480.
Kumbhakar, S., Tsionas, E., 2008. Scale and efficiency measurement using a semiparametric stochastic frontier model: evidence from the U.S. commercial banks. Empir. Econ. 34, 585–602.
Kutlu, L., Sickles, R., 2012. Estimation of market power in the presence of firm level inefficiencies. J. Econom. 168, 141–155.
Landes, W., Posner, R., 1981. Market power in antitrust cases. Harv. Law Rev. 94, 937–996.
Lerner, A., 1934. The concept of monopoly and the measurement of monopoly power. Rev. Econ. Stud. 1, 157–175.
Lindenberg, E., Ross, S., 1981. Tobin's Q ratio and industrial organization. J. Bus. 54 (1), 1–32.
McAllister, P., McManus, D., 1993. Resolving the scale efficiency puzzle in banking. J. Bank. Financ. 17, 389–405.
Mirzaei, A., Moore, T., 2014. What are the driving forces of bank competition across different income groups of countries? J. Int. Financ. Markets Instit. Money 32, 38–71.
Monti, M., 1972. Deposit, credit, and interest rate determination under alternative bank objectives. In: Szegö, G., Shell, K. (Eds.), Mathematical Methods in Investment and Finance. North-Holland, pp. 431–454.
Perloff, J., Karp, L., Golan, A., 2007. Estimating Market Power and Strategies. Cambridge University Press, Cambridge.
Scitovsky, T., 1955. Economic theory and the measurement of concentration. In: Business Concentration and Price Policy. Princeton University Press, Princeton, NJ, pp. 101–118.
Shaffer, S., 1998. Functional forms and declining average costs. J. Financ. Serv. Res. 14, 91–115.
Wheelock, D., Wilson, P., 2012. Do large banks have lower costs? New estimates of returns to scale for U.S. banks. J. Money Credit Bank. 44, 171–199.