BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Economies of scale revisited: evidence from Italian banks

by Emilia Bonaccorsi di Patti and Federica Ciocchetta

June 2020

Number 568

BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Economies of scale revisited: evidence from Italian banks

by Emilia Bonaccorsi di Patti and Federica Ciocchetta

Number 568 – June 2020

The series Occasional Papers *presents studies and documents on issues pertaining to the institutional tasks of the Bank of Italy and the Eurosystem. The* Occasional Papers *appear alongside the* Working Papers *series which are specifically aimed at providing original contributions to economic research.*

The Occasional Papers *include studies conducted within the Bank of Italy, sometimes in cooperation with the Eurosystem or other institutions. The views expressed in the studies are those of the authors and do not involve the responsibility of the institutions to which they belong.*

The series is available online at www.bancaditalia.it .

ISSN 1972-6627 (print)
ISSN 1972-6643 (online)

Printed by the Printing and Publishing Division of the Bank of Italy

ECONOMIES OF SCALE REVISITED: EVIDENCE FROM ITALIAN BANKS

by Emilia Bonaccorsi di Patti* and Federica Ciocchetta*

Abstract

This paper provides new estimates of cost scale economies for Italian banks, based on a model of bank production that takes into account a comprehensive definition of output including different categories of loans, deposits, off-balance sheet items, payment services, and brokerage and asset management activities. The output definition is more in line with the current business model of banks than previous studies since it explicitly accounts for transaction banking and IT capital. We find returns to scale in operating costs, especially for small and medium-sized institutions. For the largest institutions there is on average no statistically significant evidence of returns from scale on the cost side; however, banks falling into this latter category are quite heterogeneous in size and business model. A more extensive adoption of digital technologies in the future could expand the size range over which positive returns to scale are achievable. These results are robust to alternative input and output specifications and functional forms. An important caveat to this conclusion is that we focus solely on operating costs.

JEL Classification: D24, G21, L23.
Keywords: bank costs, scale economies, cost efficiency.
DOI: 10.32057/0.QEF.2020.568

---

* Bank of Italy, Directorate General for Economics, Statistics and Research.

Contents

1. Introduction[1]

A large literature provides estimates of returns to scale in the banking industry, obtained using data from single or multiple countries and various specifications of bank production.[2] With few exceptions, the conclusion of these analyses is that economies of scale are elusive and exhausted at small levels of output.

Nevertheless, most of the literature employs data from the 1980s and 1990s, when banks were largely providing traditional lending and deposit taking services, and does not usually take into account the expansion of banking activities into asset management and other services related to market finance. These activities have become increasingly relevant even in Europe, where most banks engage in commercial banking. In the current macroeconomic environment characterized by slow lending growth and low interest rates, banks have been diversifying their sources of revenue. While there are many studies on the effect of income diversification on risk and profitability (e.g. Stiroh, 2004), there is not much evidence on their impact on costs.[3]

Another limitation of the bulk of the literature is that it predates the digital revolution. Digital technologies are pervasive in the provision of standardized services such as payments, and the expectation is that they entail very low marginal costs and significant returns from scale (Boot, 2017). In the future, an extensive adoption of these innovations in the production of other banking outputs (e.g. lending) will probably change the cost structure of banks.

Some more recent studies document significant returns to scale in banking also for large institutions but it is difficult to disentangle the contribution of the estimation method from the impact of the change in bank production technology and in the scope of banking activities (Hughes and Mester, 2013; Wheelock and Wilson, 2012; Beccalli *et al.* 2015). Most works refer to US banking system (Wheelock and Wilson, 2015); an exception is the study by Beccalli *et al.* 2015 that estimates economies of scale for 103 European listed banks over the period 2000–2011.[4]

In this paper we present new estimates of returns to scale using data on the Italian banking industry for the period 2006-2017. Studies of cost functions for Italian banks are generally quite old and the industry has changed so much that their results are unlikely to hold today.[5] Thanks to the availability of detailed data from the supervisory reports, we can consider the full scope

[1] The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Italy or of the Eurosystem. Our thanks go to Giorgio Gobbi, for having read the various drafts of this work and for their consistently useful suggestions.

[2] Examples are Benston, Hanweck and Humphrey, 1982; Berger and Humphrey, 1994; Beccalli *et al*., 2015.

[3] One study including off-balance-sheet items and/or revenues from services finds that economies of scale exist also for large banks (Rime and Stiroh, 2003).

[4] The work by Beccalli *et al., 2015* shows that economies of scale are widespread across different size classes of banks and are more evident for small and the biggest intermediaries, especially for ones oriented towards investment banking, whereas medium-sized ones show lover economies and even diseconomies of scale. Specifically, economies of scales are larger for the smallest banks (with total assets between about 1 and 28 billion), large banks (with total assets between 182 and 553 billion) and particularly for the biggest banks (with total asset between 553 and 2590 billion). Their analysis is based on the estimation of a function for total costs, which take into account a limited number of outputs and inputs.

[5] The literature measuring scale economies for Italian banks is limited. Older studies of Italian banks find economies of scale only for small and medium-sized banks and moderate diseconomies for the largest institutions. See for example Conigliani *et al.,* 1991 and Parigi, Sestito and Viviani, 1992.

of banking outputs (traditional lending, deposits, payment services, off-balance sheet items, and brokerage and portfolio management services). Furthermore, we measure payment services by the number of transactions, which allows us to estimate directly the marginal cost of an IT intensive service and its change over time. Finally, we explicitly take into account the role of IT capital as an input, including its price in the cost function in addition to the price of labor.

Our main results are that there are significant gains from increasing the size of banks in the smallest categories and that using an extended definition of output yields larger cost economies of scale than those previously documented by other studies. We find that the elasticity of cost of payment services and services attached to deposits have declined over time, as expected, contributing to greater returns from scale in the most recent data. This result is consistent with the conjecture that technology is lowering the marginal cost of standardized products. Asset management remains, instead, a high marginal cost activity from a pure operating cost perspective, likely because it still requires specialized human capital and the development of a firm-customer relationship.

The estimated measures of returns to scale support the case for consolidation, especially for small and medium-sized banks, such as local and regional institutions. For the largest institutions there is on average no statistically significant evidence of returns from scale on the cost side; however, banks falling into this latter category are quite heterogeneous in size and business model. In general, banks that provide traditional services through physical branches tend to exhaust positive returns to scale at a lower size than others having a more diversified output mix and smaller branch networks. A more extensive adoption of digital technologies in the future could expand the size range over which positive returns to scale are achievable.

An important caveat to our findings is that we focus solely on operating costs. Indeed, risk and capital management could be both sources of returns from scale if one were to focus on overall costs, including funding costs, or on risk-adjusted profits (Hughes and Mester, 2013).

The paper is organized as follows. Section 2 describes in details the methodology used in the work and Section 3 illustrates the data and the econometric model. The results are describes in Section 4, some robustness checks are reported in Section 5 and Section 6 concludes.

2. Methodology

2.1 The cost function

Measures of scale economies are derived through the estimation of a cost function for the banking industry (Caves, Christensen and Swanson, 1981; Hunter and Timme, 1995). In very general terms, assuming that a bank i minimizes costs with respect to all inputs, there exists a total cost function for period t:

$$TC_{it} = g_t(w_{it}, y_{it}) \tag{1}$$

in which total costs TC are the sum of the quantities of inputs employed multiplied by their factor prices and are expressed as a function of the output vector y and the vector of input prices

w. The subscript t in the cost function indicates that the parameters may vary over time due to technological change and other factors. This function assumes that all input factors are variable.

If the bank does not minimize costs with respect to all inputs but only a subset of them, the use of the total cost function is not appropriate. In the latter case the bank minimizes costs with respect to some variable inputs, conditional on the level of the quasi-fixed inputs, and a variable cost function exists. Variable costs VC are given by the sum of the quantities of the variable inputs multiplied by their factor prices, and are expressed as functions of the prices of the variable inputs w, the outputs y, and the levels of the quasi-fixed inputs z:

$$VC_{it} = f_t(w_{it}, y_{it}, z_{it}) \tag{2}$$

The most common functional specification adopted in banking studies for the cost function is the second-order translog (Berger, 2003; Daglish, *et al*., 2015). The general form for the total costs function is:

$$\ln(TC_{it}) = f(\ln(w_{it}), \ln(y_{it})) + \varepsilon_{it} \tag{3}$$

where f represents the sum of terms and all cross-products and ε_{it} is the error term. In the case of a variable cost function with a quasi-fixed input, the arguments of f include also the quasi-fixed input $\ln(z_{it})$.

Starting from the above definition of the cost function, we can define two measures of returns to scale, described in details in the following subsections: Ray Scale Economies (*RSE*), Expansion Path Scale Economies (*EPSCE*).

2.2 Ray Scale Economies

Scale economies or, more precisely, returns to scale, are present when output increases at a greater rate than the one at which all input quantities are varied. The standard measure of scale economies for a multi-product firm is the Ray Scale Economies (*RSE*) index, developed by Baumol *et al.,* 1982. The *RSE* is a measure of the elasticity of costs with respect to a proportional increase in all outputs and in the case of total cost function it is defined as:

$$RSE^{TC} = \sum_i \frac{dln\, TC(w,y)}{d\, ln\, y_i} \tag{4}$$

where y_i is the ith output in the output vector.

The *RSE* index can be evaluated for each bank in the sample or for an average or representative bank in a given group, for example a size class. A value of *RSE* less than 1 indicates that costs increase proportionally less than outputs (economies of scale), while a value of *RSE* greater than 1 indicates that costs increases more than outputs (diseconomies of scales).

As shown by Caves, Christensen and Swanson, 1981 and Braeutigam and Daugherty, 1983, when the analysis is conducted with a variable cost function the measure of economies of scale has to take into account the impact of the quasi-fixed inputs on costs. The total cost function is given by the variable cost function plus the cost of the fixed factors z:

$$TC = VC(w, y, z) + r^* z \quad (5)$$

where r^* is the opportunity cost of the quasi-fixed input z. The shadow value represents the savings that would accrue to variable costs if z is increased by one unit:

$$w_z = -\frac{\delta VC(.)}{\delta z} \quad (6)$$

The equilibrium level of z is obtained when the opportunity cost equals the firm's shadow value, i.e. when:

$$-\frac{\partial VC(.)}{\partial z^*} = r^* \quad (7)$$

In equilibrium, the total (long run) and variable (short run) marginal costs are equal thus:

$$\frac{\partial TC(y, w, w_{z^*})}{\partial y} = \frac{\partial VC(y, w, z)}{\partial y} \quad (8)$$

Substituting (8) into (7) we obtain that:

$$\frac{\partial lnTC(.)}{\partial ln\, y} = \frac{\partial lnVC(y, w, z)/\partial ln\, y}{1 - \partial lnVC(y, w, z)/\partial ln\, z} \quad (9)$$

There are two methods for computing returns to scale from the variable cost function. One is to evaluate (9) using the equilibrium level of $z = z^*$ obtained using equation (7) and then substituting it into equation (5). The second is to evaluate (9) at the observed level of z (Caves, Christensen and Swanson, 1981). The second method has the advantage that it does not require knowledge of factor prices for the elements of z. In this work we follow the second method.
In the case of multiple outputs and quasi-fixed inputs, economies of scale are measured by RSE^{VC} as:

$$RSE^{VC} = \frac{\sum_i \frac{d\, lnVC(w, y, z)}{d\, ln\, y_i}}{1 - \sum_i \frac{d\, lnVC(w, y, z)}{d\, ln\, z_i}} \quad (10)$$

2.3 Expansion Path Scale Economies

One important limitation of *RSE* is that it requires all outputs to vary proportionally. As suggested by Berger, Hanweck and Humphrey, 1987, banks of different sizes tend to have different output mixes because they have access to a different opportunity set of possible outputs. They propose an alternative measure of scale economies that accounts for differences in the opportunity set across size groups, named as the index of Expansion Path Scale Economies ($EPSCE^{A,B}$).

Considering two banks, a smaller one of size A and a larger one of size B, the index $EPSCE^{A,B}$ measures the proportional change in costs as banks move from size A to size B along the expansion path observed in the industry. Therefore, *EPSCE* is the elasticity of incremental costs with respect to incremental output from A to B. For the total cost function $EPSCE^{A,B}$ is defined as

$$EPSCE^{A,B} = \left[\sum_i \frac{dln\,TC(w,y^B)}{d\,ln\,y_i} \frac{(y_i^B - y_i^A)}{y_i^B}\right]\left[\frac{(TC(w,y^B) - TC(w,y^A))}{TC(w,y^B)}\right]^{-1} \quad (11)$$

If some inputs are considered as quasi-fixed *EPSCE* can be adapted to:

$$EPSCE^{A,B} = \left\lfloor \frac{\sum_i \frac{dln\,VC(w,y^B,z^B)}{d\,ln\,y_i} \frac{(y_i^B - y_i^A)}{y_i^B}}{1 - \sum_i \frac{dln\,VC(w,y^B,z^B)}{d\,ln\,z_i}} \right\rfloor \left[\frac{(VC(w,y^B,z^B) - VC(w,y^A,z^A))}{VC(w,y^B,z^B)}\right]^{-1} \quad (12)$$

where y^A and z^A are set at the mean values of the output and fixed input bundles of banks in the small group A, y^B and z^B are set at the mean values of the output and fixed input bundles of banks in the larger group B. The input prices are the mean values for the entire industry. Values of $EPSCE^{A,B}$ less than 1 imply economies of scale, whereas values of $EPSCE^{A,B}$ greater than 1 diseconomies of scale.

Economies of scale can be evaluated for the best-practice cost function or for the average practice one, which would not take into account the efficiency levels of banks. In principle, the two methods could yield different results because inefficiencies due to factors not related to scale could bias the scale economy measures. Nevertheless, Berger and Humphrey, 1994, find that measures of scale economies for banks that are on the efficient frontier are very similar to those obtained pooling all banks. In the analysis below we compute them for the average practice cost function as most of the literature does.

3. The econometric model and data

3.1 Bank production

One of the key methodological differences across studies of economies of scale is how bank outputs, and consequently costs, are defined and modelled. Most of the literature follows either of two main approaches to bank production (see Berger and Humphrey, 1992): i) the intermediation (or asset) approach; ii) the value added approach.

Under the intermediation approach banks are considered as financial intermediaries between liability holders and receivers of bank funds (Sealey and Lindley, 1977). A bank uses labor, physical capital and financial inputs to produce financial outputs. Deposits and purchased funds are financial inputs while assets on the balance sheet, typically loans and securities, are financial outputs. Costs are given by total operating costs plus interest expenses, and outputs are measured by the amounts of dollars intermediated; prices refer to both physical (e.g. wages) and financial inputs (i.e., interest rates paid).

One major criticism to this approach is that most banks raise a substantial portion of their funds through deposits and provide liquidity, payment and safekeeping services to depositors. Depositors receive a lower interest rate on their funds because the value of these services partly compensates them. In the intermediation approach the services attached to deposits are not considered as banking products since deposits are defined a priori as an input.

The value added approach, instead, considers all liability and asset categories as having some output characteristics rather than distinguishing inputs from outputs in a mutually exclusive way. The categories yielding substantial value added are employed as important outputs, while other categories are treated as representing either unimportant outputs, intermediate products, or inputs. Berger, Hanweck and Humphrey, 1987, identify the major categories of deposits (demand, time, savings) and loans (real estate, commercial, installments) as important outputs. Purchased funds (Federal funds purchased, large CDs, foreign deposits, other liabilities for borrowed money) are treated as financial inputs because they require very little amounts of physical inputs (labor and capital). Government securities and other non-loan investments are considered as unimportant outputs, because their value added is very low. Under the value added approach costs include operating expenses only, outputs are often measured as the number of transactions completed, and only physical input prices are specified.[6]

When deposits are treated as an output the typical assumption that has to be made is that the volume of deposits is a proxy for the unpriced liquidity services produced by the bank in exchange of deposited funds. The implication is that deposits are not considered as a financial input and their cost is not included in the bank's costs of funds.

In this work we choose the value added approach for two reasons. The first reason is that we are focusing only on operating costs since our goal is to estimate scale economies originated by physical inputs (branches, IT capital and labor) so we can ignore financial costs related to bank funding and scale economies related to financial capital. The second reason is that technology is likely to have the greatest implications on the way banks provide payment services attached to deposits, so interpreting them as an output seems more appropriate in our context.[7]

[6] Most studies using the production approach use numbers of accounts as proxies for numbers of transactions, since numbers of transactions are often difficult to obtain (e.g., Berger et al., 1987, Ferrier and Lovell, 1990).

[7] In the financial firm, however, the separation between financing and technological decisions is not neutral with respect to cost minimization. For banks financial resources are essential components of the technology because banks collect deposits but apply liquidity management to transform them into loanable funds. Returns from better risk management, diversification of funding sources can affect funding costs and output mix and quality, which in turn can reduce unit costs. This aspect could be explored in future analyses.

Costs

In our main specification we analyze variable operating costs and assume that branches are a quasi-fixed input. Variable operating costs (*VOPC*) are computed as the sum of staff costs and other operating expenses (excluding depreciation and premises-related costs), commissions and fees paid by the bank related to the provision of payment services, and expenses generated by brokerage and asset management services (brokerage commissions and fees paid by the bank to service providers). We include among the costs these commissions and fees paid by the bank because they capture the use of intermediate inputs. Although they are not generated by the physical inputs directly employed by the bank, they measure the cost the bank faces when it outsources to other service providers part of the production process of the services it sells to its clients.

In our main model we exclude depreciation and premises-related costs because they refer to the quasi-fixed netput and should not enter variable costs. In an alternative model, estimated for robustness purposes, we estimate the total cost function including these expenses in total operating costs (*TOPC*). For the years 2015 and 2016 we subtract from expenses the extraordinary contribution that banks had to pay to the industry Resolution Fund because it implied a one-off increase in other operating expenses.

Inputs

The inputs included in the cost function are labor and physical capital. The latter consists of bank branches and IT capital, including hardware, software and other equipment. Labor is a variable input entering the cost function via its price, computed as average wage and benefits per employee. We include two prices, one for managers and one for other staff since their different skills imply different salaries. The data on prices are bank specific and are obtained from the national social security fund archives on wages and other compensation (INPS). While most studies use averages of labor costs at the local market level, in Italy the banking industry has a national contract fixing minimum salary levels depending on functions and worker category but banks might apply different conditions.

For the calculation of the price of IT capital we explore two alternative options. The first option is to compute the IT price as the ratio of expenses and investments related to IT to total assets. This ratio is deflated by the hedonic price index published by the Federal Reserve Bank of St. Louis because the quality of IT has changed very rapidly over time.[8] As we have no data on IT assets, we divide the expenses and investments by total assets to obtain a unit IT cost.

A second option is to follow Casolaro and Gobbi, 2007, and compute the IT price as the ratio of IT expenses to the stock of IT capital. The stock of IT capital is obtained by applying the permanent inventory method to investments in hardware, software and equipment. Investments are deflated by the hedonic price index. We apply a constant depreciation rate for the entire period, equal to 44 per cent for software and 32 per cent for hardware and equipment (Jorgenson and Stiroh, 2000).

[8] This indicator is the "Price index for private fixed investment in intellectual property products: research and development; Business: Manufacturing; Other computer and electronic product manufacturing", annual value, not seasonally adjusted.

The network of branches enters the cost function as number of branches, since we cannot account for quality differences among branches based on their location (see Hunter and Timme, 1991, Hunter and Timme, 1995).[9] For robustness purposes, in a specification where all inputs are variable, the price of premises is computed as the ratio of the expenses for premises (including depreciation) and their accounting value.

Outputs

We model banks as producing loans, deposits, payment services, services related to off-balance sheet activities, asset management and brokerage services. Other activities such as own portfolio and liquidity management are considered as intermediate inputs and do not enter the cost function.

We divide loans into different categories to account for the potential differences in the screening and monitoring costs required in their production: i) loans to households and small businesses; ii) loans to medium-sized and large businesses iii) loans to banks. The Supervisory statistics identify small businesses as sole proprietorships and limited partnerships with less than 20 employees; medium-sized and large businesses are defined as all other firms. Loans to banks include interbank loans, repurchase agreements (repos) and securities issued by banks. All loans are measured net of provisions for expected losses, to account for their credit risk. If credit quality were not taken into account, banks that have a large share of lower quality loans because they save on screening and monitoring costs would appear to be more cost efficient.

Banks hold also other financial assets; we construct a composite item “other assets” including other net loans, government bonds and other securities held (different from securities issued by banks, included in loans to banks). These categories can represent a large share of total assets for some banks and omitting them could bias the estimates of scale economies.

Liquidity services produced by the bank in exchange of deposited funds are proxied by the volume of deposits, given by the sum of checking accounts and all other deposits of customers different from banks. Since supervisory data include information on payment services, we consider them as a specific output. Payment services are measured by the number of transactions that involve checking accounts on either the credit or the debit side. We experiment with a first definition including only the number of transfers, and an extended definition including also other transactions performed through payment slips, receipts by notice (known as MAV) and direct debits through cards as their number is growing in the recent years. While the number of transactions is a direct measure of payments services, we consider the volume of deposits as capturing the provision of a distinct indivisible service of liquidity storage to clients.

Banks supply certain services that do not translate into assets nor liabilities, and excluding them would distort the estimates of the cost function. Some of these services, for example the provision of guarantees and commitments, involve origination and monitoring costs like loans, and presumably generate similar revenues (Berger and Mester, 1997). A way to proxy for them is to include the corresponding off-balance sheet items (see Clark and Siems, 2002). We measure these services by their credit equivalent reported in the Supervisory Statistics based on

[9] In principle, one could distinguish between urban and rural branches, possibly taking into account for real estate price differences.

Basel guidelines. According to the Basel guidelines, the credit equivalent has to be calculated as the amount of on balance sheet asset holdings that would result in the same amount of credit risk exposure for the bank. The use of the credit equivalent implies that most of the value added to the bank comes from the exposure to credit risk. We do not include derivatives because they have an ambiguous role as intermediate outputs and inputs, since they are mostly used for hedging and risk management purposes.

Another bank output that has become increasingly important in recent years are brokerage and asset management services. The measurement of these services is difficult because there is no information on the number of transactions performed on the account of clients. The extant literature employs the volume of revenues generated by the provision of these services, or the asset equivalent (e.g. Rime and Stiroh, 2003). In our main specification we use as a proxy the fee and commission income received by the bank for brokerage and asset management services, given that there is a specific income statement item available. For robustness purposes we use the asset equivalent amount, computed by the method proposed by Boyd and Gertler, 1994.

The asset equivalent is computed as the hypothetical volume of assets that would be required to generate the bank's flow of non-interest income. Non-interest income is capitalized using the average rate of return on assets, and the procedure is based on the following formulas.

Accounting profits (Π) are given by:

$$\Pi = IINC - IEXP - LP - NEXP + NIINC \tag{13}$$

where *IINC* is interest income, *IEXP* is interest expense, *LP* is loan provisions, *NEXP* is non-interest expense and *NIINC* is non-interest income. Assuming that *NIINC* is generated by hypothetical asset A^O and that A^O is identical to on-balance sheet assets A^B, then it is assumed that A^O and A^B have the same profitability. The asset equivalent can be obtained as:

$$AEM = A^O = A^B[NIINC/(IINC\text{-}IEXP\text{-}LP)] \tag{14}$$

In a robustness tests, we use noninterest income (*NIINC*) calculated as the difference between gross operating income and the net interest margin.

Control variables

In addition to outputs, input prices and the quasi-fixed input, we include as a control variable in the cost equation a measure of the exogenous quality of borrowers that each bank faces. If a bank is located in a risky credit market it has to employ more resources in screening and monitoring to generate the same amount of net loans than a bank operating in a safer environment. We employ the ratio of non-performing loans in the province where the bank has branches (*NPLMKT*). Since most of banks have branches in a large number of provinces we compute the branch-weighted average of the non-performing loan ratio in each province.

This variable captures not only the cross-sectional variation across banks but also the bank-specific cyclical variation in the riskiness of clients. Year dummy variables are included in the model to take into account macroeconomic conditions and other factors that can affect the banking industry such as regulatory reforms.

3.2 The Data

Our analysis is based on an unbalanced panel of banks operating in Italy in 2006-2017, excluding the two largest banking groups (Unicredit and Intesa Sanpaolo) and foreign bank branches. We exclude these two groups because they are not only much larger than other Italian banks, but also because they have specific business models. Unicredit is a multinational banking group with a significant share of activity outside Italy, and Intesa has a substantial insurance business that influences its cost structure. The accounting data are unconsolidated but all banks belonging to the same banking group are aggregated at the group level, computing pro forma balance sheets taking into account the group composition in each year. Some indicators describing the Italian banking system in the period 2006-2017 are reported in Table 1.

We let the cost function vary over time because of technological and product innovation, estimating the model separately for three different periods: i) 2006-2009 (Period 1); ii) 2010-2013 (Period 2); iii) 2014-2017 (Period 3). The cutoff dates were chosen to take into account the very different macroeconomic environments in which banks were operating; the first period ends with the beginning of the recession that affected Italy after the global financial crisis; the second period includes the sovereign debt crisis, while the third is a post-recession recovery phase, with unconventional monetary policy measures being put in place.

The cost function is estimated using data on all the sample banks; scale economies are evaluated for each bank and then averaged over banks into different size classes. The thresholds between the different classes are defined by considering the distribution of (each period average) bank assets and correspond to the following percentiles of the distribution: 50th, 75th, 90th and 95th.[10] The size cutoffs of the five classes are set to account for the skewed distribution of banks, dominated by small institutions, and are as follows: i) below €0.5 billion (very small); ii) between €0.5 and €1.45 billion (small); iii) between €1.45 and €4.3 billion (medium); iv) between €4.3 and €11.9 billion (medium-large); v) above €11.96 (large).

As a robustness check we employ another classification, based on official size classes of the Bank of Italy Statistics: class 1 includes large and medium-sized banks, class 2 small banks and class 3 minor banks. The thresholds for this definition are based on banks' earning assets and are €21.5 billion and €3.6 billion.

3.3 Specification of the cost function

We specify the cost function as a translog, the most common form adopted in banking studies using a parametric approach (Berger, 1993). Some analyses use the Fourier-flexible specification, which is a translog augmented with Fourier trigonometric terms, which provides greater flexibility than the translog. A significant drawback of the Fourier-flexible specification is that the number of parameters that need to be estimated is much larger, which requires a large number of observations. Berger and Mester, 1997 find that the improvement obtained using the

[10] The definition of thresholds are evaluated over all the period 2006-2017, so the class thresholds are fixed over time. A bank can change its size class from a period to the other as its (average) size changes, due for instance to mergers and incorporation or an increase of assets. A description of the banks' transition between different size classes is reported in Table 2.

Fourier-flexible is not significant from an economic point of view, even if the additional parameters are jointly statistically different from zero, so they suggest employing the translog.

Our main specification is therefore based on the following functional form (Model I):

$$\ln\left(\frac{VOPC}{w_1}\right) = \alpha + \sum_{i=2}^{M}\beta_i \ln\left(\frac{w_i}{w_1}\right) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=2}^{M}\beta_{ij}\ln\left(\frac{w_i}{w_1}\right)\ln\left(\frac{w_j}{w_1}\right) + \sum_{j=1}^{K}\gamma_j \ln(y_j) + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}\gamma_{ij}\ln(y_i)\ln(y_j) + \sum_{j=1}^{N}\delta_j \ln(z_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\delta_{ij}\ln(z_i)\ln(z_j) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=1}^{K}\tau_{ij}\ln\left(\frac{w_i}{w_1}\right)\ln(y_j) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=1}^{N}\varphi_{ij}\ln\left(\frac{w_i}{w_1}\right)\ln(z_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{K}\vartheta_{ij}\ln(z_i)\ln(y_j) + \sum_{g=1}^{G} Controls_g + \sum_{t=1}^{Tp} T_t + \varepsilon_{it} \tag{15}$$

The dependent variable *VOPC* stands for variable costs; the w_i for i = 1,…, M=3 are the prices of variable inputs (labor price for employees, labor price for managers, and price of IT capital); the y_j for j=1,…, K=8 are the outputs and include loans to households and small businesses (y_1), loans to medium-large businesses (y_2), loans to banks (y_3), other assets (y_4), credit equivalent for contingent liabilities and commitments (y_5), deposits (y_6), payment transactions (y_7), brokerage and asset management services (y_8); z is the number of branches, the only quasi-fixed input (N=1). All quantity variables are deflated by the CPI index (base 2006). The vector of control variables contains the market NPL ratio (*NPLMKT*) and the T_t dummy variables for years that capture neutral technical change or other exogenous factors, such as changes in regulation.

A description of all the variables is reported in Table 3. Costs and prices are normalized by one of the input prices (w_1) to impose linear price homogeneity.[11] The standard symmetry restrictions are imposed ($\beta_{ij} = \beta_{ji}$, $\gamma_{km} = \gamma_{mk}$, $\delta_{rs} = \delta_{sr}$).

As in most of the recent literature, we do not estimate the cost function jointly with share equations embodying Shephard's Lemma restrictions because this would impose the undesirable assumption of no allocative inefficiencies.[12]

The translog cost function is estimated by OLS over the whole sample of banks separately for each period, allowing the coefficient to vary across periods to reflect changes in technology, regulation and market environment. The error terms are clustered at the bank level.

[11] A doubling of all prices exactly doubles costs in the cost function.

[12] From Shephard's Lemma cost share equations can be derived for each input by taking the derivative of the cost function with respect to the price of each input, s_j = dln(TC)/dlnP_j (Greene, 2008); this yields a system of N optimal cost share equations, one for each of the inputs used in the production function. The cost shares must sum to one, as each is the proportion of the total costs spent on that input.

4. Results

4.1 Elasticities of costs to outputs

We do not report the coefficient estimates for the cost function for the sake of brevity but we derive the variable cost elasticities with respect to each output to assess the consistency with economic intuition. Each elasticity is defined as the percentage change of variable costs from a one percentage change in each output, keeping the other outputs constant.[13] The cost elasticities, calculated with respect to each output and across periods, are reported in Table 4.
The elasticities exhibit significant variability across outputs and over time. All of the elasticities, except for the one with respect to the credit equivalent of off-balance sheet items, are positive indicating that an increase in the output leads to an increase in variable costs as expected. The negative elasticity of off-balance sheet items likely depends on multicollinearity between the services captured by this product and the volume of lending to firms.

The magnitude of the elasticities are generally consistent with our priors. The highest cost elasticity is the one of brokerage and asset management services, and has increased over time. This product is costly to produce because it requires relatively specialized labor but has also become an important source of income for banks.

The elasticities of the different categories of loans can be compared since the underlying outputs are measured with the same metric. The cost elasticity of net loans to medium and large firms is the largest in magnitude, and has increased over time, possibly reflecting an increase in credit risk. The elasticity of net loans to households and small firms has instead declined, and is very small in the last period. A possible explanation is that banks have increased loans to households at the expense of riskier loans to small firms, resulting in a drop of the share of assets invested in the latter (the median decreased from about 39 per cent in the first period to 29 per cent in the last one).

We note that the elasticity of costs to deposits and payment services have both declined to very low values, signaling that the marginal impact of these products on bank costs is very small, consistent with an increased role of digital technology.

4.2 Ray Scale Economies

The estimated values of the *RSE* index for the main model (Model 1) are reported in Table 5. We evaluate Equation (10) for each bank using the estimated coefficients, setting the exogenous variables at the observed levels for that bank. We then compute the average of the scale indices for the corresponding period and size class. Statistical significance is evaluated against the null hypothesis that the *RSE* is equal or greater than 1, estimating standard errors by the delta method.

[13] In a standard log-log model, the elasticity of the dependent variable with respect to an independent variable is given by the coefficient in the equation model. In the case of a translog function, a number of nonlinear terms are added to the model and these should be taken into account in the derivation of the elasticity. The elasticity is evaluated as the average elasticity for banks in each size class; similar results are obtained when the elasticity is evaluated at the average value of the exogenous variables for each bank size class.

The value of the *RSEs* support the existence of increasing returns to scale for all but the largest size class of banks in the two earlier periods (Table 5, Panel A). In the most recent years increasing returns to scale are statistically significant for classes 3-5. For the smallest two size categories the *RSE* are 0.85 and 0.76, respectively. The median values of the *RSEs* are consistent with these results (Table 5, panel B) and similar findings are obtained if we employ different definitions of size classes (Table 6).

We note that the individual bank values of the *RSEs* are somewhat heterogeneous within each size class in each period, especially for the relatively larger institutions (Figure 1). Moreover, banks with different business models and of quite different size fall into the two largest classes. Some have a high deposit to assets ratio and at the same time a large branch network, which suggest that they have only partially exploited technology in service provision. This could have occurred either because of supply or because of demand constraints. Other banks have a more diversified output mix and a limited number of branches given their size.

Focusing on the time dimension, our findings suggest that on average banks in size classes 1 and 2 have generally exhausted increasing returns to scale in operating costs in the last period, given their current business model. New opportunities could arise if they were to adopt technology in a transformative way. Banks in classes 4 and 5 are instead facing greater opportunities in recent years than in the past because they are still far from the optimal size for their current business model. For banks in class 3 the *RSE* has remained essentially unchanged.

4.3 Elasticities of costs to branches

We employ the cost equation to derive the elasticity of variable costs with respect to the number of branches. We report the average elasticity for each period and size class.

The results, shown in Panel C of Table 5, indicate that an increase in the number of branches by 1 percent increases variable costs by 0.18 percent in the first period, by 0.14 percent in the second period, and by 0.18 percent in the last period. However, there are interesting differences between classes and over time. First, the magnitude of the elasticity for medium-sized and large banks is lower than the one for small banks in each period. This indicates that an increase or decrease in the number of branches has a greater impact on variable operating costs for small banks than for the others. Second, comparing the magnitudes estimated for period 3 with those for period 1, we observe that the elasticity declines for large banks and increases for small ones.

The drop in the elasticity for large banks can be explained by the sharp reduction in the number of branches of the largest banking groups. The marginal benefit of reducing branches is declining as banks proceed with reorganizing their branch networks. For small banks the slight increase in the elasticities in recent years suggests that these intermediaries have ample room to improve in terms of optimizing branches.

4.4 Expansion Path Scale Economies

Table 7 reports the estimates of *EPSCE* for each size class, based on the parameters estimated for Model 1. We find that there are expansion path scale economies for most bank size classes, and for all periods. The notable exception is if a bank were to increase in size from class 2 to class 1 (the largest size class in our sample), for which the *EPSCE* is greater than one. This finding is consistent with the conclusion derived from the *RSEs* that the largest institutions in the sample do not enjoy increasing returns from scale in operating costs.

An intuition about the magnitude of *EPSCE* can be gained by comparing the representative (average) banks in each size class. Table 8 (Panels A and B) shows the output mix and the number of branches, scaled by total assets, for the average bank in each size class and the differences of these values between two consequent classes (i+1 and i, for i = 1, 2, 3, 4) for the latest data (2014-2017).

For a bank increasing its size from class 5 to class 4 the *EPSCE* is 0.92. The number indicates that a representative bank growing along the path described by the change in output mix between the representative bank in class 5 and the one in class 4 would face a less than proportional increase in operating costs. Considering that the average bank in size class 5 is 235 million in total assets while the one in size class 4 is 881 million, the size increase would be about 3.7 times. The value of *EPSCE*, which is based on a linear approximation, suggests that operating costs would increase by about 3.4 times.

The *EPSCE* is smaller than the one of the *RSE* because the latter index does not take into account the changes in output mix that a bank growing in size would undertake as a result of competitive pressure (see Berger and Humphrey, 1991). In this case of an increase from size 5 to size 4 the bank would not differ substantially in terms of output mix, it would have a larger share of loans to other firms, an output that has a relatively high cost elasticity.

If a banks moves from class 4 to 3, it increases from 0.8 to 2.5 billion in total assets. The main change in the output mix is a further increase in corporate lending, and a reduction in the share of some of the low marginal cost activities, which explains why *EPSCE* (0.95) is lower than the corresponding *RSE* index. Further increases in size yield even lower returns from scale because (*EPSCE* is 0.98) because brokerage services would tend to increase.[14] A bank that moved from class 2 to class 1, instead, would experience an increase in assets by almost eight times, and would face variable cost diseconomies. The output mix would be characterized by a larger share of assets invested in corporate loans and loans to other banks, and a reduction in payment services and deposits.

The estimation of the *EPSCEs* is derived from the variable operating costs function, considering branches as fixed at the average level of banks in each size class. The figures reported in Table 8 show that the ratio of branches to total assets (million) declines as banks increase in size. For the largest institutions one branch manages on average 100 million of assets, more than twice as much the volume managed at small banks. The reduction in branches, per asset managed, would be another source of cost savings on top of the benefits in terms of variable operating costs.

[14] The average bank in size 2 is about 2.92 times in terms of total assets than the average bank in class 3. The implied increase in costs would be about 2.86 times, based on the linear approximation.

5. Robustness

5.1 Cost function specification

We consider a number of alternative specifications of the cost function, to check for robustness of results to different proxies of outputs or prices, and different aggregations of output types.

The first model is that of a cost function with no quasi-fixed inputs letting branches vary and including the price of premises. The results on scale economies from our main specification are confirmed (Table 9).

The results of a large number of models based on alternative definitions of outputs and input prices are shown in the Appendix. In Model II we measure loans gross of provisions (Table A1). Model III employs the asset equivalent definition of brokerage and asset management services, based on the hypothesis that the return on these outputs is equal to the return of lending and other investment (Table A2). In Model IV instead these services are measured by non-interest income (*NIINC*, Table A3). In Model V we proxy payment services with an extended definition, which includes transfers, payment slips, receipts by notice, direct debts though cards (Table A4). In Model VI we compute the ICT using the permanent inventory method ICT following Casolaro and Gobbi 2007 instead of using the ratio of ICT expenses and investments to assets from accounting data (Table A5). Overall, results are very similar to those obtained with the main model in all of these alternative specifications of the output vector.

As a further robustness check, we estimated two specifications based on a smaller number of outputs to reduce the risk of running into multicollinearity among the variables. In the first of these two specifications loans to large firms are aggregated with the credit equivalent of off-balance sheet items, and other assets include a number of wholesale assets (loans to banks, other loans, securities) for a total of six outputs (Model VII). In the second one we drop other assets (Model VIII). The *RSEs* again are very similar to those of our main model (Tables A6 and A7).

The conclusion from these tests is that the measures of economies of scale are very stable, and are not affected by changes in the definitions of outputs and inputs, nor by our choice of considering branches as a quasi-fixed netput, nor depend on the functional form imposed on the data.

5.2 Functional form

Although the translog is widely employed in the estimation of cost functions it is unable to fit a monotonically decreasing average cost function (Shaffer, 1998). The implication of this property is that it could fail to recognize decreasing returns from scale in the entire range of the data. For robustness purposes with respect to this risk we consider a translog augmented with the terms $1/\ln(y_i)$ and $1/\ln(y_i)^2$ for each output, as suggested by Shaffer, 1998. These terms introduce greater flexibility in the functional form and do not require more complex nonparametric approaches (McAllister and McManus, 1993)).

We consider the following functional form:

$$\ln\left(\frac{VOPC}{w_1}\right) = \alpha + \sum_{i=2}^{M} \beta_i \ln\left(\frac{w_i}{w_1}\right) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=2}^{M} \beta_{ij} \ln\left(\frac{w_i}{w_1}\right)\ln\left(\frac{w_j}{w_1}\right) + \sum_{j=1}^{K} \gamma_j \ln(y_j) + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} \gamma_{ij} \ln(y_i)\ln(y_j) + \sum_{j=1}^{N} \delta_j \ln(z_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \delta_{ij} \ln(z_i)\ln(z_j) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=1}^{K} \tau_{ij} \ln\left(\frac{w_i}{w_1}\right)\ln(y_j) + \frac{1}{2}\sum_{i=2}^{M}\sum_{j=1}^{N} \varphi_{ij} \ln\left(\frac{w_i}{w_1}\right)\ln(z_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{K} \vartheta_{ij} \ln(z_i)\ln(y_j) + \sum_{g=1}^{G} Controls_g + \sum_{t=1}^{Tp} T_t + \sum_{j=1}^{K} \gamma_j\, 1/\ln(y_j) + \sum_{i=1}^{K} \theta_{ii}\, 1/(\ln(y_i)\ln(y_i)) + \varepsilon_{it} \tag{16}$$

The variables are as in equation (17). We then compute the *RSEs* as (10). The results, reported in Table 10, do not reject our main finding that there are economies of scale for medium-sized and small banks, and some diseconomies for larger banks.

In other tests we employ a translog augmented with the terms $1/y_i$ for each output, and an extension of the hyperbolically augmented Cobb-Douglas (HACD), in which some terms of the form $1/y_i$ and some interaction terms between outputs and prices are added to a standard Cobb-Douglas. Results, available upon request, show that average values of the *RSE* are similar to our main results.

6. Conclusions

We estimate measures of returns to scale in operating costs for the Italian banking industry and find that there are significant benefits in terms of costs from increasing the size of small and medium-sized banks. Benefits for larger institutions are more difficult to identify since banks in this category are more diverse in size and business model. Our results show that the marginal cost of standardized services such as payments and deposits has declined in recent years, likely because of an extensive adoption of new technologies in their production. Lending to corporate borrowers and asset management services remain high marginal cost products, possibly because their provision entails significant human capital and tailoring.
An important caveat to our findings is that we focus solely on operating costs. Indeed, risk and capital management could be both sources of returns from scale if one were to focus on overall costs, including funding costs, or on risk-adjusted profits.

Tables and figures

Table 1: Descriptive statistics

Panel A: Number of banks and branches, staff and banks' total assets

Year	Number of banks	Number of branches	Staff headcount	Total assets (billions)
2006	575	32,191	333,873	2,618
2007	571	32,882	332,229	2,639
2008	562	33,907	330,585	3,097
2009	553	33,676	316,547	3,186
2010	542	33,146	312,984	3,204
2011	540	33,110	296,928	3,373
2012	528	32,491	302,721	3,583
2013	512	31,232	298,376	3,437
2014	495	29,861	288,204	3,336
2015	484	29,475	286,752	3,219
2016	461	28,833	288,400	3,270
2017	407	27,198	275,896	3,117

Panel B: Balance sheet indicators

Year	ROE	ROA	Operating Costs/Total Assets	Cost-to-income ratio	Staff Costs /Operating Costs	Net Fee Income/ Gross Income
2006	12.1	0.9	1.7	54.9	55.5	26.1
2007	8.2	0.8	1.7	58.9	54.4	26.6
2008	4.1	0.4	1.6	63.8	53.2	26.3
2009	2.7	0.2	1.5	63.3	51.8	27.8
2010	3.0	0.3	1.4	65.1	53.0	31.3
2011	-7.8	-0.7	1.4	66.1	52.7	30.8
2012	-1.1	-0.1	1.3	63.1	52.9	29.7
2013	-8.3	-0.6	1.2	59.2	51.6	29.7
2014	-3.5	-0.3	1.3	60.2	51.5	31.0
2015	1.4	0.1	1.5	61.8	48.9	31.6
2016	-7.7	-0.6	1.5	76.5	49.8	36.2
2017	5.4	0.4	1.5	71.8	50.7	37.4

Source: Supervisory reports on unconsolidated data for individual banks. Note: percentage values, weighted averages of data referring to all resident banks (including Unicredit and Intesa SanPaolo); foreign branches are excluded.

Table 2: Matrix of bank size migration
Panel A: From period 1 to period 2

	size class	Period 2: 2010-2013					
		1	2	3	4	5	out
Period 1: 2006-2009	1	**19**					
	2	4	**14**				1
	3		7	**34**			8
	4			16	**61**		9
	5			1	62	**271**	44
	in		3	1	4	13	8

Panel B: From period 2 to period 3

	size class	Period 3: 2014-2017					
		1	2	3	4	5	out
Period 2: 2010-2013	1	**17**					6
	2	3	**14**	1			6
	3		5	**40**	2		5
	4			14	**84**	1	28
	5				34	**176**	74
	in	1		2	6	6	57

Source: Supervisory data. Note: Tables report the possible variation of banks' size classes between two periods. The elements on the diagonal indicate the number of banks which maintain the same classification in the two periods. The elements in the lower parts of the tables indicate the number of banks which move from class *k* in the period *t* to the class *k+1* in the period *t+1* whereas in the upper part from class *k* in period t to class *k-1* in the period *t+1*. "In" indicates new banks in the period *t+1* whereas out the banks that are present in the period t but they are not in the period *t+1*.

Table 3: Descriptive statistics for variables employed in the cost function estimation

The table shows averages by size class and period, with size classes defined as follows: i) below €0.5 billion (very small); ii) between €0.5 and €1.45 billion (small); iii) between €1.45 and €4.3 billion (medium); iv) between €4.3 and €11.9 billion (medium-large); v) above €11.96 (large).

		Period 1: 2006-2009					**Period 2: 2010-2013**					**Period 3: 2014-2017**				
VARIABLES\ SIZE CLASSES		**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Banks	Number of banks	19	19	49	86	378	23	24	52	127	284	21	19	57	126	182
Assets	Billion euros (deflated)	68.2	7.3	2.6	0.9	0.2	69.4	8.2	2.6	1.0	0.2	60.9	7.6	2.6	0.9	0.3
Costs																
VOPC	Staff costs, other operating expenses (different from branches) and commission fees.	1,029	163	54	18	5.4	900	119	48	15	5.1	920	146	42.5	16.0	5.4
Variable input prices																
w1	Price of labor for employees (thousands).	45	38	42	41	40	45	42	44	43	43	43	41	45	43	43
w2	Price of labor for managers (thousands).	210	175	181	173	131	188	183	185	167	132	177	173	189	169	131
w3	Price of IT as the ratio of IT investments plus expenses over assets (mean, %).	0.32	0.32	0.29	0.27	0.35	0.28	0.20	0.27	0.23	0.36	0.34	0.23	0.25	0.26	0.35
Variable output quantities																
y1	Loans to households and small business (million).	11,700	2,040	647	292	82	13,041	1,710	602	270	80	12,855	1,586	560	255	71
y2	Loans to other firms (million).	15,516	1,825	654	255	46	14,737	1,534	614	213	43	11,737	1,000	544	165	34

y3	Loans to banks, bank securities and repos (million).	10,240	846	306	64	21	12,084	835	288	84	24	9,910	698	222	80	28
y4	Other assets: loans and securities different from the above categories (million).	24,460	1,856	768	204	60	28,531	3,194	855	268	70	26,073	3,989	1,169	382	101
y5	Commitments, letters of credit and other off-balance-sheet items excluding derivatives, measured using Basel' credit equivalent (million).	5,450	413	135	55	10	4,230	205	80	27	7.5	4,928	140	62	21	5.7
y6	Payment services (number of transactions).	3.25	0.70	0.13	0.06	0.01	9.25	1.14	0.38	0.14	0.05	9.09	1.38	0.46	0.19	0.06
y7	Deposits (million).	19,381	2,663	1,023	382	109	23,799	3,384	1,080	380	110	25,668	4,208	1,251	472	139
y8	Brokerage and asset management services, proxied by fee and commission income (million).	548	89	30.4	6.9	1.8	545..6	69.9	29.5	6.0	1.9	561.9	99.9	21.0	6.9	2.3
Fixed input quantities																
z1	Number of branches.	724	109	40	18	6	662	95	35	16	5	609	72	33	16	5
Controls																
NPLMKT	Market-average of the nonperforming loan ratio (percentage points).	7.5	7.5	7.5	7.4	11.7	12.3	11.0	10.4	12.2	13.5	19.1	18.2	18.1	19.0	20.2

Table 4: Elasticities of variable costs to outputs – Model I

For each output the elasticity is calculated as the average of elasticities evaluated at the output levels observed for each bank in each period. The total elasticity is computed as the average of the sum of the y1-y8 elasticities.

	Period 1: 2006 - 2009	**Period 2: 2010 - 2013**	**Period 3: 2014 - 2017**
Loans to households and small firms (y1)	0.03	0.08	0.02
Loans to other firms (y2)	0.06	0.10	0.11
Loans to banks (y3)	0.02	0.06	0.05
Other assets (y4)	0.02	0.07	0.06
Off-balance sheet items (y5)	-0.01	-0.02	-0.03
Payment services (y6)	0.07	0.05	0.002
Deposits (y7)	0.13	0.05	0.006
Brokerage & asset management services (y8)	0.40	0.44	0.49
Total	0.73	0.74	0.70

Table 5: Ray Scale Economies and Cost Elasticities by size class and period - Model I

Input: labor (both employees and executives) and IT; Quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other assets, credit equivalent**,** deposits**,** payment services and commission income as proxy of asset management services. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period. Panel B reports median values of ray scale indexes by size class and period. Panel C shows the average elasticities of costs to branches by class size and period. Size classes are identified by the following percentiles of distribution of banks' (average) assets over the whole period: 50, 75, 90 and 95.

Panel A: Average RSE by class and period

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.99	1.01	1.08
Class 2	0.95 (***)	0.94 (***)	0.99
Class 3	0.89 (***)	0.89 (***)	0.91 (***)
Class 4	0.88 (***)	0.85 (***)	0.85 (***)
Class 5	0.84 (***)	0.78 (***)	0.76 (***)

Panel B: Median value of RSE by class and period

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	1.00	1.03	1.06
Class 2	0.97	0.94	0.99
Class 3	0.91	0.90	0.91
Class 4	0.89	0.85	0.85
Class 5	0.85	0.80	0.77

Panel C: Elasticity of costs to branches

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.18	0.05	0.06
Class 2	0.16	0.07	0.11
Class 3	0.14	0.09	0.14
Class 4	0.18	0.12	0.17
Class 5	0.19	0.17	0.21
(weighted average)	0.18	0.14	0.18

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table 6: Ray Scale Economies by size class – Model I, alternative size classes

Input: labor (both employees and executives) and IT; Quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other assets, credit equivalent**,** deposits**,** payment services and commission income as proxy of asset management services. Ray scales are calculated according to Equation 10. In both panels values are the average of ray scale indexes for each size class and period. Size classes in Panel A are identified by the following percentiles of the distribution of average assets over the three different periods: 1st quartile, median value and 3rd quartile (size classification 2 or quartile-based classification). The classification in Panel B is defined in terms of the size class used in Bank of Italy's main publications (size classification 3): 1 includes large and medium banks, 2 small banks and 3 minor banks.

Panel A: Alternative class 2

Size class 2	**Period 1: 2006 - 2009**	**Period 2: 2010 - 2013**	**Period 3: 2014 - 2017**
Quartiles			
Class 1	0.97 (**)	0.97 (**)	1.04
Class 2	0.89 (***)	0.89 (***)	0.91 (***)
Class 3	0.88 (***)	0.85 (***)	0.85 (***)
Class 4	0.84 (***)	0.79 (***)	0.75 (***)

Panel B: Alternative classes 3

Official size classes (size class 3)	**Period 1:** 2006 - 2009	**Period 2:** 2010 - 2013	**Period 3:** 2014 - 2017
Class 1	0.98	0.98	1.06
Class 2	0.95 (***)	0.98 (**)	1.04
Class 3	0.91 (***)	0.88 (***)	0.87 (***)

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table 7: Expansion Path Scale Economies (EPSCE) – Model I

Input: labor (both employees and executives) and IT; Quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other assets, credit equivalent**,** deposits**,** payment services and commission income as proxy of asset management services. Expansion Path Scale indicators are calculated according to Equation 11. Size classes are identified by the following percentiles of the distribution of average assets over the three different periods: 50, 75, 90 and 95.

Expansion Path from class A to B	**Period 1: 2006 - 2009**	**Period 2: 2010 - 2013**	**Period 3: 2014 - 2017**
From class 2 to 1	0.94	0.99	1.03
From class 3 to 2	0.93	0.96	0.98
From class 4 to 3	0.91	0.96	0.95
From class 5 to 4	0.92	0.93	0.92

Table 8: Output mix for bank size classes
(assets in millions, other variables in percentage values)
Panel A: Output mix for each size class (in terms of assets)

Output/size class	**class 1**	**class 2**	**class 3**	**class 4**	**class 5**
Loans to households and small firms (y1)	18.3	18.8	19.3	24.9	26.2
Loans to other firms (y2)	16.7	11.8	18.8	16.1	12.6
Loans to banks (y3)	14.1	8.3	7.7	7.8	10.1
Other assets (y4)	19.2	28.6	27.8	26.1	26.5
Credit equivalent (y5)	7.0	1.7	2.1	2.1	2.1
Payment services (y6)	15.0	19.0	18.4	21.6	25.6
Deposits (y7)	36.5	49.8	43.2	46.2	51.1
Brokerage & asset management services (y8)	0.8	1.2	0.7	0.7	0.8
Branches (z1)	1.0	1.0	1.4	1.8	2.1
Assets (millions)	60,573	7,273	2,494	881	235

Panel B: Delta between output mix moving from class *i+1* to class *i*

	From class 2 to class 1	From class 3 to class 2	From class 4 to class 3	From class 5 to class 4
Loans to households and small firms (y1)	-0.5	-0.5	-5.6	-1.3
Loans to other firms (y2)	4.8	-7.0	2.7	3.5
Loans to banks (y3)	5.8	0.6	-0.1	-2.3
Other assets (y4)	-9.4	0.8	1.8	-0.4
Credit equivalent (y5)	5.3	-0.5	0.0	0.0
Payment services (y6)	-4.0	0.5	-3.1	-4.0
Deposits (y7)	-13.3	6.6	-3.0	-5.0
Brokerage & asset management services (y8)	-0.4	0.5	0.1	0.2
Branches (z1)	0.0	-0.4	-0.5	-0.3

Note: The tables refer to the period 2014-2017. For each size class, the average bank is considered. All the variable values are defined in terms of total assets, with the exception of assets.

Table 9: Robustness – Ray scale economies based on Total Costs model

Dependent variable: Total costs; Input prices: labor (employees and executives), IT, cost related to premises and branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other investments, credit equivalent, deposits, payment services and commission fees as a proxy of asset management services. Ray scales are calculated according to Equation 4. We report values the average of ray scale indexes for each size class and period.

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.99	0.99	1.04
Class 2	0.93 (***)	0.91 (***)	0.90 (***)
Class 3	0.91 (***)	0.87 (***)	0.88 (***)
Class 4	0.91 (***)	0.87 (***)	0.87 (***)
Class 5	0.88 (***)	0.80 (***)	0.79 (***)

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

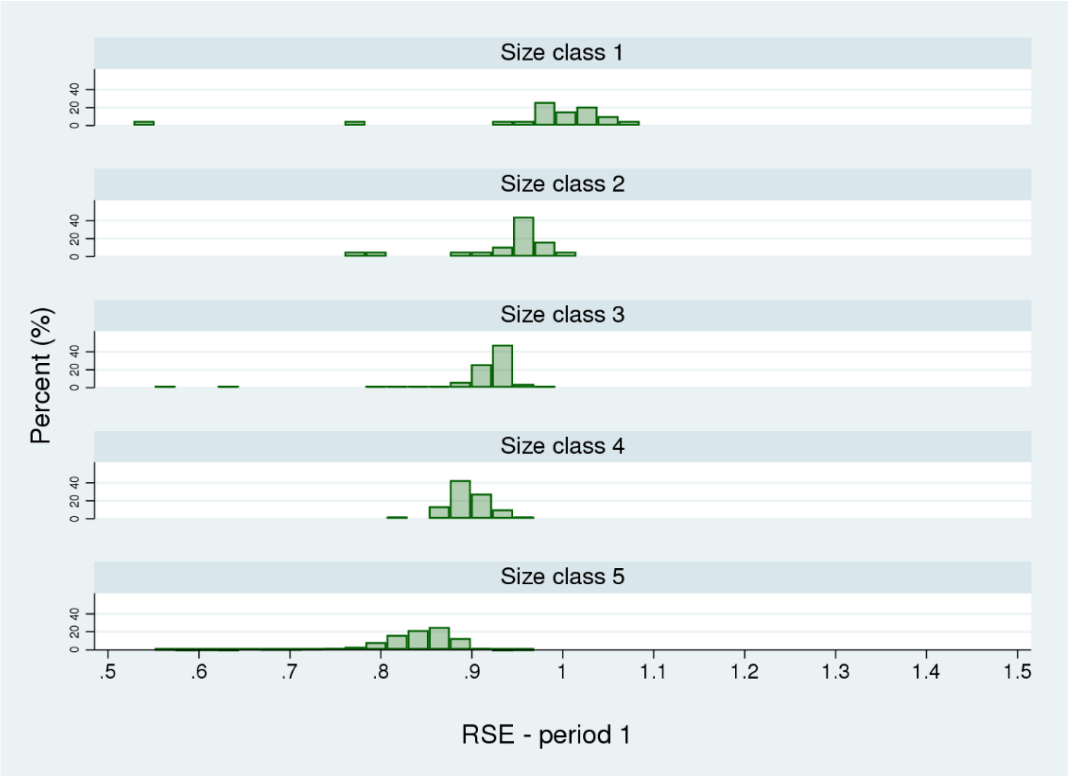
Table 10: Robustness – Alternative functional form

The table shows results based on a translog augmented with the terms 1/ln(yi) and $1/\ln(yi)^2$ for each output. The specification of the cost function is the one of Model I. Input: labor (both employees and executives) and IT; Quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other assets, credit equivalent**,** deposits**,** payment services and commission income as proxy of asset management services. Ray scales are calculated according to Equation 10. We report the average of ray scale indexes for each size class and period. Size classes are identified by the following percentiles of the whole period distribution of banks' (average) assets: 50, 75, 90 and 95.

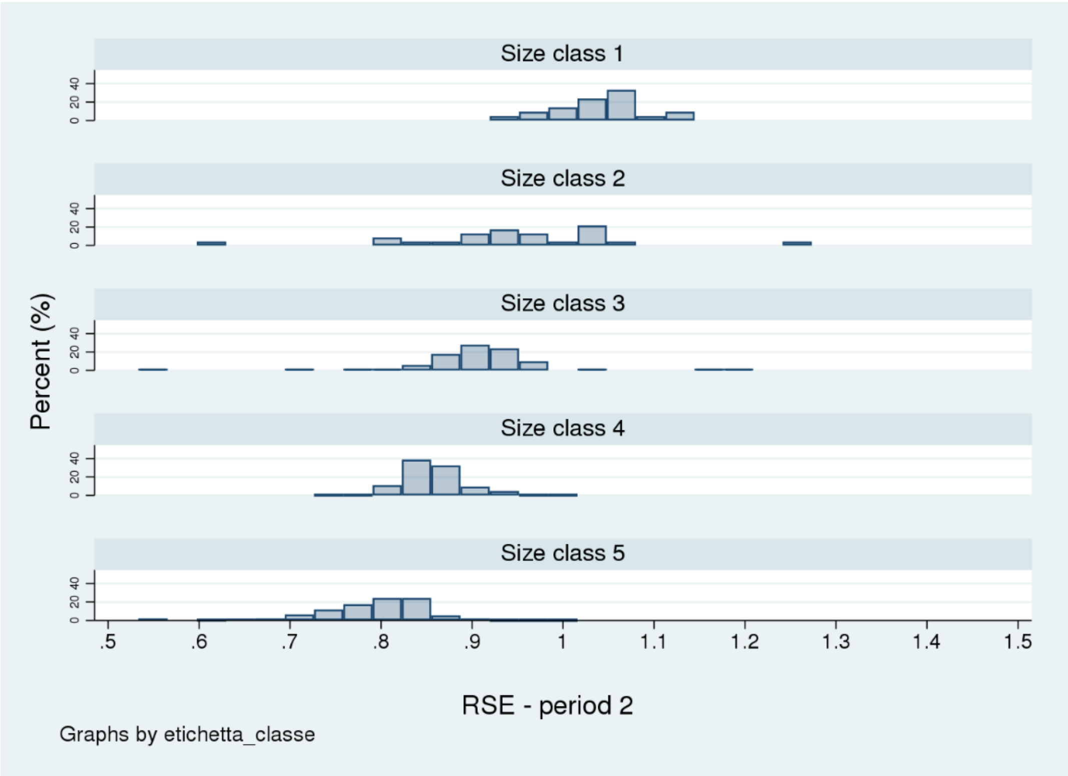
Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.92	0.96	1.05
Class 2	0.93	0.95	1.03
Class 3	0.92	0.95	0.92
Class 4	0.95	0.90	0.87
Class 5	0.84	0.79	0.79

Figure 1: Distribution of RSE - Model I

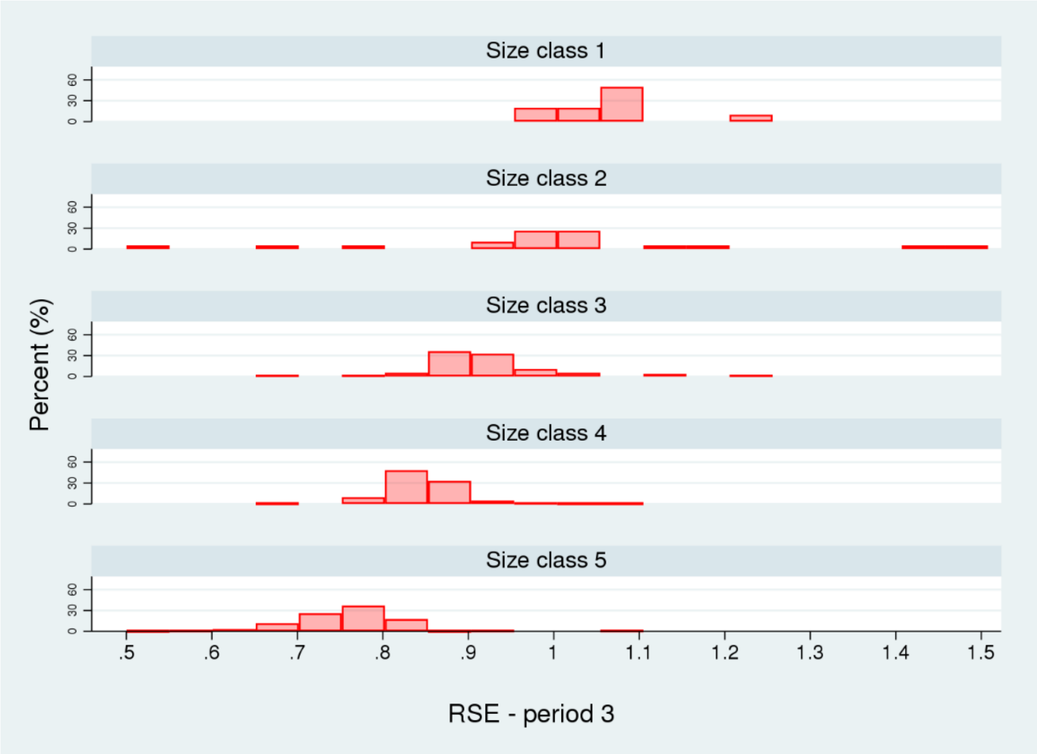
Period 1



Period 2



Period 3



Note: The height of the bars are scaled so that the sum of their heights equals 100 (percent value).

APPENDIX

Table A1: Robustness - Loans gross of provisions (Model II)

Input: labor (both employees and executives) and IT; Quasi-fixed input: branches; Output: gross loans to households and SMEs, gross loans to other firms, loans to banks, other assets, credit equivalent, deposits, payment services and commission income as proxy of asset management services. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.99 (*)	1.00	1.08
Class 2	0.95 (***)	0.93 (***)	0.99
Class 3	0.89 (***)	0.88 (***)	0.92 (***)
Class 4	0.88 (***)	0.85 (***)	0.86 (***)
Class 5	0.84 (***)	0.78 (***)	0.76 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.15	0.03	0.02
Class 2	0.14	0.05	0.07
Class 3	0.13	0.08	0.09
Class 4	0.16	0.11	0.13
Class 5	0.19	0.17	0.19
Average over period	0.17	0.13	0.14

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A2: Robustness – Asset Equivalent for Brokerage and Asset Management Services (Model III)

Input: labor (both employees and executives) and IT; quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other investments, credit equivalent**,** deposits, payment services and asset equivalent of brokerage and asset management revenues.
Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 -2009	Period 2: 2010 -2013	Period 3: 2014 - 2017
Class 1	0.97	1.00	1.12
Class 2	0.89 (*)	0.92 (*)	0.96 (***)
Class 3	0.88 (***)	0.88 (***)	0.95 (***)
Class 4	0.88 (***)	0.86 (***)	0.88 (***)
Class 5	0.85 (***)	0.83 (***)	0.78 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 -2009	Period 2: 2010 -2013	Period 3: 2014 - 2017
Class 1	0.18	0.19	0.07
Class 2	0.19	0.18	0.10
Class 3	0.17	0.15	0.12
Class 4	0.21	0.16	0.14
Class 5	0.24	0.17	0.17
Average over period	0.23	0.16	0.15

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A3: Robustness – Non-Interest Income for Brokerage and Asset Management Services (Model IV)

Input: labor (both employees and executives) and IT; quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other investments, credit equivalent**,** deposits, payment services and non-interest income. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.98	1.03	1.13
Class 2	0.96 (*)	0.97 (*)	1.02
Class 3	0.91 (***)	0.92 (***)	0.96 (***)
Class 4	0.90 (***)	0.88 (***)	0.91 (***)
Class 5	0.87 (***)	0.82 (***)	0.82 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.15	0.11	0.07
Class 2	0.15	0.13	0.08
Class 3	0.13	0.11	0.10
Class 4	0.18	0.13	0.13
Class 5	0.21	0.17	0.16
Average over period	0.19	0.14	0.12

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A4: Robustness - Extended definition of payment services (Model V)

Cost: variable costs; Input: labor (employees and executives) and IT; quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms, loans to banks, other investments, credit equivalent, deposits, payment services (extended definitions including transfers, payment slips, receipts by notice, direct debts through cards) and commission income**.** Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 -2009	Period 2: 2010 -2013	Period 3: 2014 - 2017
Class 1	0.98	1.01	1.08
Class 2	0.94 (***)	0.94 (***)	1.00
Class 3	0.90 (***)	0.89 (***)	0.91 (***)
Class 4	0.89 (***)	0.85 (***)	0.85 (***)
Class 5	0.84 (***)	0.78 (***)	0.76 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 -2009	Period 2: 2010 -2013	Period 3: 2014 - 2017
Class 1	0.12	0.04	0.10
Class 2	0.12	0.07	0.13
Class 3	0.12	0.09	0.15
Class 4	0.15	0.11	0.18
Class 5	0.18	0.17	0.21
Average over period	0.17	0.14	0.18

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A5: Robustness - IT price by the permanent inventory method (Model VI)

Input prices: labor (employees and executives) and IT; quasi-fixed input: branches; output: loans to households and SMEs, loans to other firms, loans to banks, other investments, credit equivalent, deposits, payment services, and commission income as proxy for brokerage and asset management services. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 - 2009	Period 2: 2010 - 2013	Period 3: 2014 - 2017
Class 1	0.99	1.02	1.10
Class 2	0.94 (***)	0.94 (***)	0.99
Class 3	0.90 (***)	0.89 (***)	0.91 (***)
Class 4	0.89 (***)	0.84 (***)	0.84 (***)
Class 5	0.84 (***)	0.77 (***)	0.74 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 -2009	Period 2: 2010 -2013	Period 3: 2014 - 2017
Class 1	0.13	0.03	0.07
Class 2	0.12	0.06	0.04
Class 3	0.12	0.08	0.13
Class 4	0.15	0.10	0.16
Class 5	0.18	0.17	0.21
Average over period	0.17	0.16	0.17

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A6: Robustness – Six outputs (Model VII)

Cost: variable costs; Input: labor (employees and executives) and IT; quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms (including credit equivalent), other assets, deposits, payment services and commission income. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray scale economies

Size class	Period 1: 2006 - 2009	Period 2: 2010 – 2013	Period 3: 2014 - 2017
Class 1	0.97 (*)	0.97	1.02
Class 2	0.97 (*)	0.93 (*)	0.96
Class 3	0.95 (***)	0.91 (***)	0.93 (***)
Class 4	0.94 (***)	0.90 (***)	0.89 (***)
Class 5	0.91 (***)	0.86 (***)	0.84 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 - 2009	Period 2: 2010 – 2013	Period 3: 2014 - 2017
Class 1	0.11	-0.04	0.10
Class 2	0.11	0.03	0.11
Class 3	0.14	0.07	0.11
Class 4	0.17	0.14	0.14
Class 5	0.20	0.21	0.15
Average over period	0.18	0.15	0.13

Note: In the calculation of the *RSE* means, we do not consider few outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

Table A7: Robustness – Five outputs (Model VIII)

Cost: variable costs; Input: labor (employees and executives) and IT; quasi-fixed input: branches; Output: loans to households and SMEs, loans to other firms (including credit equivalent), exposure to banks, deposits, payment services and commission income. Ray scales are calculated according to Equation 10. Panel A reports averages of ray scale indexes by size class and period; Panel B reports average elasticities of costs to branches by size class and period.

Panel A: Ray Scale Economy

Size class	Period 1: 2006 - 2009	Period 2: 2010 – 2013	Period 3: 2014 - 2017
Class 1	0.98 (*)	0.98	1.05
Class 2	0.98	0.94 (***)	1.00
Class 3	0.93 (***)	0.90 (***)	0.91 (***)
Class 4	0.91 (***)	0.89 (***)	0.87 (***)
Class 5	0.88 (***)	0.84 (***)	0.81 (***)

Panel B: Elasticity of costs to branches

Size class	Period 1: 2006 - 2009	Period 2: 2010 – 2013	Period 3: 2014 - 2017
Class 1	0.16	-0.012	0.13
Class 2	0.12	0.05	0.13
Class 3	0.15	0.07	0.13
Class 4	0.18	0.14	0.16
Class 5	0.20	0.19	0.16
Average over period	0.19	0.15	0.15

Note: In the calculation of the *RSE* means, we exclude a small number of outliers with extreme or negative values. We apply one-side t-test of *RSE* less than one against the null hypothesis of *RSE* = 1. The symbols *, **, *** indicate that the unitary *RSE* hypothesis is rejected, with a p-value greater than 0.99, 0.95, 0.90, respectively.

References

Baumol, W.J., J.C. Panzar and R.D. Willig, 1982. "*Contestable Markets and the Theory of Industry Structure"*, New York, Harcourt Brace Jovanovich.

Beccalli, E., M. Anolli and G. Borello, 2015. "Are European banks too big? Evidence on economies of scale", *Journal of Banking and Finance*, Elsevier, Vol. 58(C): 232-246.

Benston, G.J., G.A. Hanweck and D.B. Humphrey, 1982. "Scale Economies in Banking: a Restructuring and Reassessment", *Journal of Money, Credit and Banking*, Vol. 4: 435-456.

Berger, A.N., 1993. "Distribution Free Approach: Estimates of Efficiency in the US banking Industry and tests of the Standard Distributional Assumptions", *The Journal of Productivity Analysis*, Vol. 4: 261-292.

Berger, A.N., 2003. "The Economic Effects of Technological Progress: Evidence from the Banking Industry", *Journal of Money Credit and Banking*, Vol. 35: 141-76.

Berger, A.N., G.A. Hanweck and D.B. Humphrey, 1987. "Competitive Viability in Banking", *Journal of Monetary Economics*, Vol. 20: 501-520.

Berger, A.N. and D.B. Humphrey, 1991. "The dominance of inefficiencies over scale and product mix economies in banking", *Journal of Monetary Economics*, Vol. 28 (1): 117-148.

Berger, A.N. and D.B. Humphrey, 1992. "Measurement and efficiency issues in commercial banking", in Zvi Griliches (ed.), *Output measurement in the service sectors, NBER, Studies in Income and Wealth*, University of Chicago Press, Vol. 56: 245-79.

Berger, A.N. and D.B. Humphrey, 1994. "Bank Scale Economies, Mergers, Concentration, and Efficiency: the U.S. Experience", Financial Institutions Center, Wharton University, Working Paper N. 25.

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

Boyd, J.H. and M. Gertler, 1994. "Are Banks Dead? Or Are the Reports Greatly Exaggerated?", Federal Reserve Bank of Minneapolis Quarterly Review, Vol. 18 (3): 2-23.

Boot, Arnoud W.A., 2017. "The Future of Banking: From Scale & Scope Economies to Fintech", *European Economy*, Vol. 2.

Braeutigam, R.R. and A.F. Daughety, 1983. "On the Estimation of Returns to Scale Using Variable Cost Functions", *Economic Letters*, Vol. 11: 25-31.

Casolaro L. and G. Gobbi, 2007. "Information, Technology and Productivity Changes in the Banking Industry", *Economic Notes*.

Caves, D.W., L.R. Christensen and J.A. Swanson, 1981. "Productivity Growth, Scale Economies, and Capacity Utilization in US Railroads, 1955-74", *American Economic Review*, Vol. 71 (5): 994-1002.

Clark, J.A. and T.F. Siems, 2002. "X-Efficiency in Banking: Looking Beyond the Balance Sheet", *Journal of Money, Credit and Banking,* Vol. 34 (4): 987-1013.

Conigliani, C., R. De Bonis, G. Motta, and G. Parigi, 1991. "Economie di Scala e di Diversificazione nel Sistema Bancario Italiano", Temi di discussione, N. 150, Bank of Italy.

Daglish, T, O. Robertson, D. Tripe and L. Weill, 2015. "Translog Cost Function Estimation: Banking Efficiency," Working Paper Series 4180, Victoria University of Wellington, The New Zealand Institute for the Study of Competition and Regulation.

Ferrier, G.D. and C.A.K. Lovell, 1990. "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence", Journal of Econometrics, Vol. 46 229-245.

Greene, W. H., 2008. "Econometric analysis". Granite Hill Publishers.

Hughes, J.P. and L. J. Mester, 2013. "Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function", Journal of Financial Intermediation Vol. 22(4): 559-585.

Hunter, W.C., and Timme S.G., 1991. "Some evidence on the Impact of Quasi-Fixed Inputs on Bank Scale Economy Estimates", *Economic Review*, May/June 1991.

Hunter, W.C., and Timme S.G., 1995. "Core Deposits and Physical Capital: A Reexamination of Bank Scale Economies and Efficiency with Quasi-Fixed Inputs", *Journal of Money, Credit and Banking*, Vol. 27(1): 165-185.

Jorgenson, D. and K.J. Stiroh, 2000. "Raising the Speed Limit: U.S. Economic Growth in the Information Age *Brookings Papers on Economic Activity* Vol. 1: 125-212.

McAllister, P.H. and D. McManus, 1993. "Resolving the scale efficiency puzzle in banking", *Journal of Banking & Finance,* Vol. 17 (2-3): 389-405.

Parigi, G., P. Sestito and U. Viviani, 1992. "Economie di scala e di diversificazione nell'industria bancaria: il ruolo dell'eterogeneità tra imprese", Temi di discussione, Bank of Italy, N. 174.

Rime, B. and K.J. Stiroh, 2003. "The Performance of Universal Banks: Evidence from Switzerland", *Journal of Banking and Finance,* Vol. 27: 2121-50.

Sealey, C.W. and J.T. Lindley, 1977. "Inputs, Outputs and a Theory of Production and Cost at Depository Financial Institutions", *Journal of Finance,* Vol. 32 (4): 1251-66.

Stiroh, K. J., 2004. "Diversification in Banking: Is Non-interest Income the Answer?", *Journal of Money, Credit, and Banking,* Vol. 36 (5), October.

Wheelock, D. and P. Wilson, 2012. "Do large banks have lower costs? New estimates of returns to scale for US banks," *Journal of Money, Credit, and Banking,* Vol. 44: 171-99.

Wheelock, D. and P. Wilson, 2015. "The Evolution of Scale Economies in U.S. Banking", FRB St. Louis Working Paper N. 2015-21.