



The productivity growth of euro area banks

Ivan Huljak¹ · Reiner Martin² · Diego Moccero³

Accepted: 4 April 2022 / Published online: 7 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

We study the performance of the banking system in the Eurozone over the period 2006–2017 as measured by total factor productivity growth (TFPG) and its components. We find that Total Factor Productivity growth for the median euro area bank decreased from around 2.6% in 2007 to below 1.7% in 2017, driven mainly by a decline in technical efficiency. In addition, we control for unobserved heterogeneity across banks and disentangle persistent and time-varying inefficiency in the banking sector. This modelling choice is important to avoid distorted and biased inefficiency estimates. We find that cost efficiency in the euro area banking sector amounted to around 84% on average over the 2006 to 2017 period. The largest part of bank inefficiency is persistent, suggesting that structural long-term factors (such as location, client structure, macroeconomic environment, regulation, etc.) play a bigger role than time-varying factors.

Keywords Euro area banking sector · total factor productivity · cost-efficiency frontier · panel data · time-varying inefficiency · persistent inefficiency

JEL Classification C23 · D24 · G21

1 Introduction

The analysis of productivity (i.e., the increase in output unrelated to the amount of inputs employed) and its components in the euro area banking sector is important for several reasons. First, as banks are the largest providers of credit to companies and households in the euro area, boosting efficiency in the banking sector helps to ensure lower lending rates and higher lending volumes. Schlüter et al. (2012) and Shamshur and Weill (2019) find for European countries that more efficient banks have lower loan mark-ups and provide credit at lower interest rates, fostering access to credit and benefiting borrowers and the economy.

Second, a more efficient banking sector improves the transmission of monetary policy. As such, monetary policy relaxation (for example) is associated with a larger increase in loans when the banking sector is more efficient (Jonas and King 2008).¹ Moreover, Schlüter et al. (2012) and Havranek et al. (2016) find that more efficient banks tend to smooth loan interest rates for their clients when market rates change, hence contributing to lower volatility in bank lending interest rates. Third, Dissanayake and Wu (2021) find that more efficient banking systems experienced lower stock market price volatility during the Covid-19 pandemic. And fourth, Fiordelisi et al. (2011) find that more inefficient banks tend to be riskier, in the sense that they have a higher default frequency, hence contributing to increase risks to financial stability.

Boosting productivity in the euro area banking sector is also critical because banks have been struggling to maintain profitability since the Global Financial Crisis (GFC), and although profitability has recovered somewhat since then, it

✉ Diego Moccero
Diego.Moccero@ecb.int

¹ Croatian National Bank, Trg hrvatskih velikana 3, 10000 Zagreb, Republic of Croatia

² Joint Vienna Institute, Mariahilferstrasse 97, A-1060 Vienna, Austria

³ European Central Bank, Sonnemannstrasse 22, 60314 Frankfurt am Main, Germany

¹ They find both theoretically and empirically that the marginal cost curve of producing loans is flatter for more efficient banks, resulting in a flatter supply curve. As a result, following a relaxation in monetary policy, the supply curve shifts to the right more strongly for efficient banks, resulting in a larger increase in bank lending.

remains below the long-run cost of capital, estimated to be in the range of 8–10% (EBA 2018). Several factors acted as a drag on bank profitability. The low-interest-rate environment resulted in lower lending margins, and low-income diversification have put pressure on income growth. Moreover, the increase observed in legacy assets in several countries has also dragged on profitability, not only because they must be recognised and provisioned for, but also because they tie scarce capital without providing returns, absorbing operational capacity and incurring legal and administrative costs. A high cost-base is also to blame, which has been associated with extensive branch networks and overbanking. Returning banks to sustainable rates of profitability is important for several reasons. First, it ensures that the sector remains resilient, because profits are the first line of defence against losses from credit impairment, which increase in times of crisis. Second, low profitability prospects translate into low bank valuations, which hinder the ability of banks to raise equity capital. And third, banks with poor structural profitability are likely to face higher funding costs and may also be tempted to take on more risk (Babihuga and Spaltro 2014).

Against the need to improve the health of the euro area banking sector, we study the performance of the banking system in the Eurozone over the period 2006–2017 as measured by total factor productivity growth (TFPG) and its components. The first one is overall technical efficiency (*i.e.*, the minimisation of costs to generate a certain amount of output, for given input prices). The second one, technological progress, captures the decline (or increase) in total costs over time, for a given amount of output and input prices. The third component is the equity effect and captures the impact of equity ratios on total costs. The last one is the scale effect and captures the impact of output growth on average costs. All these components can be computed based on cost-efficiency frontier analysis. As a result, it is possible to isolate changes in productivity growth related to moving closer to the cost frontier, shifts in the frontier itself, cost reductions due to higher levels of equity and moving to a different part of the cost-efficiency frontier, respectively.

Our contribution to the empirical literature is threefold. First, we quantify the four drivers of Total Factor Productivity growth for a large sample of euro area banks over a period that covers by and large the aftermaths of the global financial crisis and the euro area sovereign debt crisis. Focusing on this period is important because the euro area banking sector has been struggling to achieve structural profitability due to constraints in its capacity to increase revenue in a low-interest-rate environment, impediments to implement business models adjustments that allow for an increase in income other than interest revenue and downsize costs, and excess capacity. As a result, improving productivity could help re-build profits in the banking sector.

Second, to the knowledge of the authors, this is the first study that controls for unobserved bank heterogeneity and disentangles persistent and time-varying inefficiency in the euro area banking sector. Our modelling choice is important because failing to control for unobserved bank heterogeneity and to distinguish between persistent and time-varying inefficiency is likely to yield distorted and biased inefficiency estimates (Colombi et al. 2014; Kumbhakar et al. 2014; Badunenko and Kumbhakar 2017). And third, we test for the impact of the global financial crisis and the sovereign debt crisis on total costs.

We employ stochastic frontier analysis (SFA), the most widely used parametric method for measuring firm-specific cost-efficiency. The sample is based on a panel of commercial, cooperative and savings banks from 17 euro area countries over the period from 2006 to 2017. We estimate a trans-log cost function to capture banks' relative ability to convert inputs (financial capital, labour and other aggregate inputs) into outputs (loans and investments), while minimising costs. The methodology is appealing because it allows the estimation of overall cost-efficiency, technological progress and the equity and scale effects within the same econometric framework.

We find that total factor productivity growth (TFPG) in the euro area banking sector decreased from around 2.6% in 2007 to below 1.7% in 2017. The largest component of TFPG (in terms of absolute size) is technological progress. The contribution of this component increased during the last ten years, from 3% in 2007 to 3.4% in 2017. Technical efficiency change is the second-largest component and the main culprit for the decline observed in TFPG over the last decade. For the median bank, the negative contribution of technical efficiency change increased from -0.8% in 2006 to -1.9% in 2017. The scale and equity effect have a minor impact on TFPG. All in all, the decline observed in TFPG in the euro area banking sector is undesirable given the need to enhance the profitability of euro area banks.

Looking at the components of Total Factor Productivity, we find that overall cost efficiency for the median euro area bank amounted to around 84% on average over the period from 2006 to 2017, implying that a bank operating on the frontier could produce the same level of output with 84% of the current costs. Moreover, we find that the largest part of bank inefficiency is persistent, suggesting that structural long-term factors (such as location, client structure, macroeconomic environment, regulation, etc.) play a bigger role than time-varying factors. We also find that the shadow cost of equity computed from our model delivers comparable results to those derived from the Capital Asset Pricing Model (CAPM) and point to the fact that the reward for being a better-capitalised institution increased in times of financial stress. Turning to economies of scale, we find that they tend to be larger for smaller institutions, although the largest institutions also exhibit economies of scale.

The rest of the paper is organised as follows. Section 2 presents the literature review, with a focus on studies for European countries. Section 3 decomposes TFPG into its main components. The econometric model disentangling persistent and time-varying inefficiency in the euro area banking sector is presented in Section 4. The data are presented in Section 5 and the econometric results in Section 6. In Section 7, we present empirical results for the various components of TFPG, namely technical efficiency, technological progress and the equity and the scale effect. Section 8 concludes.

2 Literature review

There are several studies estimating cost functions for banks in Europe based on stochastic frontier analysis. Most papers have tended to focus on specific components of Total Factor Productivity. Altunbas et al. (1999) employ a stochastic cost frontier estimation technique to study the impact of technological progress (decomposed into pure, scale augmenting and non-neutral) on the costs of European banks over the period from 1989 to 1996. They find that the rate of reduction in costs due to technological progress increased between 1989 and 1996 and that larger banks benefited more. Beccalli et al. (2015) investigate the presence of economies of scale for European listed banks over 2000–2011 and find that economies of scale are widespread across different size classes of banks and are especially large for the largest banks. Maudos et al. (2002) and Bos and Schmiedel (2007) analyse cost and profit efficiency in a sample of European banks. The former find that profit efficiency is lower than cost efficiency and the latter find evidence of a common technology, which is supportive of a single and integrated European banking market. Maudos et al. (2002) and Nițoi and Spulbar (2015) study the determinants of cost efficiency, including bank size, specialisation, bank-specific characteristics (liquidity, capitalisation, etc.), features of the markets in which they operate and macroeconomic variables. Stronger GDP growth, macroeconomic stability, less market concentration, lower network density and healthier institutions in terms of liquidity and capitalisation are positively correlated with cost efficiency.

Other studies look at a larger number of drivers of Total Factor Productivity. Altunbas et al. (2001) and Boucinha et al. (2013) used a cost function to estimate three components of Total Factor Productivity, namely cost efficiency, return to scales and technological progress in the German (1989 to 1996) and the Portuguese (1992 to 2006) banking sectors, respectively. Differentiating by ownership types (state-owned, mutual and private institutions), Altunbas et al. (2001) find that all three bank ownerships benefit from widespread economies of scale, but inefficiency measures indicate that public and mutual banks have slight cost

advantages over their private-sector competitors. Boucinha et al. (2013) found that scale economies have also contributed to boost productivity in the Portuguese banking sector. Casu et al. (2016) focus on a larger number of drivers of TFPG in nine euro area countries, namely scale efficiency change, technical change, changes due to environmental factors, changes in allocative efficiency and changes in cost efficiency. They estimate a parametric eurozone-level, meta frontier for commercial banks and find that technological spillovers have led to progression toward the best technology. Accounting for several components of Total Factor Productivity allows for a comparison of their most important drivers and the literature suggests that technological progress is the component that made the most important contribution to cost reduction in the European banking system.

Other studies link inefficiency estimates to other banking variables or have estimated the impact of structural changes on productivity and efficiency. For example, Altunbas et al. (2007) and Fiordelisi et al. (2011) apply stochastic frontier analysis to estimate the efficiency of European banks and subsequently use time series econometric techniques to assess the inter-temporal relationship between bank efficiency, capital and risk over the period 1992–2000 and between 1995–2007, respectively. The two papers find opposite results regarding the relationship among these variables. Casu et al. (2016) find that the introduction of the single currency in 1999 appears to have enhanced bank productivity. Psillaki and Mamatzakis (2017) investigate the effects of financial regulations and structural reforms on the cost efficiency of the banking industries of 10 Central and Eastern European (CEE) countries for the period 2004–2009 and find that the extent to which government borrowing crowds-out private borrowing results in lower efficiency. They also find that better-capitalized banks are more cost efficient. Bonin et al. (2005) use a stochastic frontier and conclude that privatization by itself is not sufficient to increase bank efficiency. However, they find that foreign-owned banks are more cost-efficient than other banks and that they also provide better service, particularly if they have a strategic foreign owner.

While making relevant contributions to the understanding of bank productivity and its components in European countries, an important caveat of most of the studies mentioned above is that they assume that all inefficiency is time varying and do not control for unobserved bank heterogeneity. From the studies mentioned above, only Maudos et al. (2002) account for unobserved heterogeneity. However, their model fails to distinguish between time-varying and persistent inefficiency, confounding the latter into heterogeneity. This is unfortunate because one of the biggest advantages of panel data models is their superior ability to take heterogeneity into account. Moreover, we expect persistent inefficiency to play an

important role in the banking sector because of the large sunk costs needed to operate in this business. These observations are relevant because an important line of research focused on panel data econometric models that capture and separate both unobserved bank heterogeneity and persistent inefficiency. Some recent studies for non-European countries disentangle these two components of banking sector inefficiency. Badunenko and Kumbhakar (2017) used the methodology to conclude, among others, that state banks in India were able to improve their cost efficiency, while Indian private banks were lagging. Fungačova et al. (2020), employing a sample of 166 Chinese banks during the 2008–2015 period, find similar contributions of persistent and time-varying inefficiency. To the knowledge of the authors, ours is the first study that disentangles persistent and time-varying inefficiency in the euro area banking sector, while controlling for unobserved bank heterogeneity.

3 Decomposing total factor productivity growth (TFPG)

We employ a *component-based approach* to productivity and decompose the Total Factor Productivity Growth (TFPG) of euro area banks into its main components (Balk 2001; Kumbhakar et al. 2015). The technology is specified by the dual cost function, under cost minimisation behavioural assumptions: $C = C(w, z, y, t)e^\varphi$, where $\varphi \geq 0$ is input-oriented technical inefficiency, w are inputs, z is a quasi-fixed input, y are outputs and t is time. Differentiating the cost function totally one obtains:

$$\begin{aligned} \dot{C} &= \sum_j \frac{\partial \ln C}{\partial \ln w_j} \dot{w}_j + \frac{\partial \ln C}{\partial \ln z} \dot{z} + \frac{\partial \ln C}{\partial \ln y} \dot{y} + \frac{\partial \ln C}{\partial \ln t} + \frac{\partial \varphi}{\partial t} \\ &= \sum_j S_j \dot{w}_j - \dot{Z} + \frac{1}{RTS} \dot{y} - TPROG - TEC \end{aligned} \tag{1}$$

where $S_j = \frac{\partial \ln C}{\partial \ln w_j}$, $\dot{Z} = -\frac{\partial \ln C}{\partial \ln z} \dot{z}$, $\frac{1}{RTS} = \frac{\partial \ln C}{\partial \ln y}$, $TPROG = -\frac{\partial \ln C}{\partial \ln t}$ and $TEC = -\frac{\partial \varphi}{\partial t}$. Differentiating total costs $C = w'x$, where x are input prices, gives: $\dot{C} = \sum_j S_j(\dot{w} + \dot{x})$. Equating this equation and Eq. (1), one obtains the equation for Total Factor Productivity Growth (TFPG):

$$TFPG = \dot{y}(1 - RTS^{-1}) + \dot{Z} + TPROG + TEC$$

In this equation, TFPG is Total Factor Productivity Growth, TEC is the rate of growth in technical efficiency, TPROG is technological progress, RTS is returns to scale, and \dot{Z} is the impact of equity on total costs. All the components are measured in percent.

Technical efficiency (TEC) measures how close a bank is to the technology frontier. Put differently, it is the relative ability of a bank to convert inputs (financial capital, labour and other aggregate inputs) into outputs (loans and

investments), while minimising costs.² The most efficient bank is the one that has the lowest cost while generating a certain amount of output, for given input prices. Therefore, the efficiency results are relative to the best practise bank, rather than absolute.

Technological progress (TPROG) captures the decline (or increase) in total costs over time, for a given amount of output and input prices. It results from a shift in the production technology. According to Baltagi and Griffin (1988) and Kumbhakar and Heshmanti (1996), technological progress can be divided into three components. The first one is “pure technological progress” and captures only on the impact of time on total costs. The second is called “scale-augmenting technological progress” and captures the change in the sensitivity of total costs with respect to time, as output changes. The third component is called “non-neutral technological progress” and reflects the changes in the sensitivity of total costs to time, as input prices change.

The third and fourth components of TFPG are the so-called “equity effect” and “scale effect”. The latter captures the importance of operating at the optimal scale and the fact that a bank can increase its productivity by changing the scale of its operations (Kumbhakar et al. 2014 and 2015) and the former captures the impact of the shadow cost of equity and changes in the equity ratio on Total Factor Productivity Growth. It measures the impact of a change in equity on bank costs in a particular year.

The four components of TFPG can be computed based on stochastic frontier analysis (SFA), the most widely used parametric method for measuring firm specific cost-efficiency.³ The assumption behind this methodology is that the distance from the frontier is not entirely under the influence of the bank due to both random error and the functional form of the cost function.⁴ The methodology is appealing because it allows the estimation of technical

² Farrell (1957) pioneered the work on firm inefficiency and defined it as a waste of resources, measured by the ratio between minimal (derived from a benchmark firm) and observed production costs. This work provided the ground for the future development of frontier methods.

³ The cost-efficiency frontier can also be computed using non-parametric approaches, based on linear programming. This approach works well with small samples and does not require a priori assumptions on the functional form of the best practise frontier. However, non-parametric techniques do not allow for random error in the model, making the efficiency scores sensitive to changes in the definition of inputs and outputs. Parametric approaches to cost-efficiency frontier analysis developed into three directions: stochastic frontier analysis (SFA) introduced by Aigner et al. (1977) and Battese and Corra (1977), distribution free approach (DFA) and thick frontier approach (TFA). In this paper we focus on SFA because TFA does not allow for computing bank specific efficiency while DFA does not compute year by year efficiency scores.

⁴ Therefore, SFA is sometimes referred to as composed error, since the part of the cost that cannot be explained by outputs and input prices is divided into an idiosyncratic random error and inefficiency.

efficiency, technological progress and the equity and scale effects from the same econometric model.

4 The econometric model

Traditional panel data econometric models often do not separate individual heterogeneity from unobserved, persistent inefficiency, as the model will tend to confound persistent inefficiency with heterogeneity, captured by a single, bank-specific effect in the model.⁵ As a result, Greene (2005) developed the “true” random or fixed effect model in which the author separates firm-specific effects and persistent inefficiency. Other models treat all firm effects as persistent inefficiency and include another component to capture time-varying technical inefficiency (e.g., Kumbhakar 1991; Kumbhakar and Heshmati 1995; and Kumbhakar and Hjalmarsson 1995). Still another strand of the literature has considered all inefficiency as time varying, with or without controlling for unobserved bank heterogeneity (Maudos et al. 2002, Casu et al. 2016, Bos and Schmiedel 2007, and Boucinha et al. (2013), respectively). However, none of these two specifications may be fully satisfactory, as inefficiency is likely to be partly persistent and partly time varying. In fact, persistent inefficiency is likely to be important in the banking industry because there are large sunk costs associated with starting a bank and it requires several years of deposit base formation to succeed in the business. Moreover, it tends to be costly to restructure a bank (downsize the number of staff, merge the bank with another institution, etc.).

Because inefficiency scores are sensitive to how inefficiency is modelled and because persistent inefficiency is expected to be large in the banking industry (as mentioned before), we employ the generalized “true” random-effects (GTRE) model proposed by Colombi et al. (2014), Kumbhakar et al. (2014) and Filippini and Greene (2016).⁶ On top of capturing time-varying inefficiency, this model decomposes the bank-specific effect into a random, bank-specific effect (capturing unobserved heterogeneity à la Greene 2005) and persistent technical inefficiency. In total, this model decomposes the error term of the stochastic cost function into four components, namely: *i*) time-varying inefficiency; *ii*) persistent (time-invariant) inefficiency; *iii*) a bank-specific effect, capturing latent heterogeneity across banks; and *iv*) a pure random noise component (Greene 2005).

Therefore, the stochastic cost function can be written as follows:

$$\ln TC_{it} = \alpha_0 + \ln TC(y_{it}, w_{it}; \beta) + \psi_i + v_{it}^+ + \eta_i^+ + u_{it} \quad (2)$$

where α_0 is a constant, i refers to the cross-sectional unit and t refers to time, TC_{it} represents total costs, $TC(y_{it}, w_{it}; \beta)$ is a function of outputs and input prices, y_{it} are outputs produced by bank i at time t , w_{it} are input prices, β is a vector of parameters, ψ_i and $\eta_i^+ > 0$ are a bank-specific effect and persistent (time-invariant) inefficiency, respectively. $v_{it}^+ > 0$ and u_{it} are time-varying inefficiency and the random error, respectively. Finally, \ln denotes the natural logarithm.

The function $TC(y_{it}, w_{it}; \beta)$ represents the cost frontier while the sum of the constant (including the bank-specific effect), the function $TC(y_{it}, w_{it}; \beta)$ and the idiosyncratic error represent the stochastic frontier. The difference between total costs and the stochastic frontier is the measure of cost inefficiency.

Equation (2) can be rewritten as:

$$\ln TC_{it} = \alpha_{0^*} + \ln TC(y_{it}, w_{it}; \beta) + \alpha_i + \epsilon_{it} \quad (3)$$

where $\alpha_{0^*} = \alpha_0 + E(\eta_i^+) + E(v_{it}^+)$, $\alpha_i = \psi_i + \eta_i^+ - E(\eta_i^+)$ and $\epsilon_{it} = u_{it} + v_{it}^+ - E(v_{it}^+)$.

To operationalise the calculation of the efficiency scores, we employ the pseudo maximum likelihood estimation recommended by Kumbhakar et al. (2014). By contrast, the approaches employed by Colombi et al. (2014) and Filippini and Greene (2016) are more complex to estimate.⁷ In particular, we estimate the model following the three-stage procedure: *i*) We run the standard random-effects panel regression model to estimate β and to predict the values of α_i and ϵ_{it} ; *ii*) We estimate the time-varying technical inefficiency, v_{it}^+ using the predicted values of ϵ_{it} from the first step. In particular, for $\epsilon_{it} = u_{it} + v_{it}^+ - E(v_{it}^+)$, we apply standard Stochastic Frontier Analysis (SFA) using Maximum Likelihood by assuming that u_{it} is *i.i.d.* $N(0, \sigma_u^2)$ and v_{it}^+ is *i.i.d.* $N^+(0, \sigma_v^2)$ *iii*) We apply a similar approach as in the second step for $\alpha_i = \psi_i + \eta_i^+ - E(\eta_i^+)$. In particular, we apply standard SFA cross-sectionally, assuming that ψ_i is *i.i.d.* $N(0, \sigma_\psi^2)$ and η_i^+ is *i.i.d.* $N^+(0, \sigma_\eta^2)$ in order to obtain estimates of the persistent technical inefficiency component η_i^+ ; *iv*) Finally, overall technical inefficiency is computed as the product of persistent technical inefficiency and time-varying technical inefficiency.

We employ a trans-log cost function for $TC(y_{it}, w_{it}; \beta)$ with three inputs and two outputs, while including both a linear and a quadratic time trend and the bank capital ratio to

⁵ Berger (1993 and 1995) show that bank specific effects tend to confound differences in bank size with inefficiency.

⁶ The three papers differ in their estimation procedure. See the discussion below.

⁷ The former uses full information maximum likelihood and may provide more efficient coefficient estimators, but it has been judged to be impracticable in general. The latter is an intermediate case between Kumbhakar et al. (2014) and Colombi et al. (2014). It uses maximum simulated likelihood applied to two skew-normal density functions.

capture technological progress and risks considerations, respectively. In our framework, banks produce loans and other earning assets (outputs), while utilising labour, physical capital and financial funds (inputs).⁸ As a result, Eq. (3) can be written as follows:

$$\begin{aligned}
 \ln TC_{i,t} = & \alpha_0 + \phi_c + \sum_{h=1}^2 \alpha_h \ln y_{h,i,t} + \sum_{j=1}^3 \beta_j \ln w_{j,i,t} \\
 & + \tau_0 \ln E_{i,t} + t_1 T \\
 & + \frac{1}{2} \left[\sum_{h=1}^2 \sum_{k=1}^2 \delta_{hk} \ln y_{h,i,t} \ln y_{k,i,t} + \sum_{k=1}^3 \sum_{j=1}^3 \gamma_{kj} \ln w_{k,i,t} \ln w_{j,i,t} \right. \\
 & \left. + \varphi_1 \ln E_{i,t} \ln E_{i,t} + t_{11} T^2 \right] \\
 & + \sum_{h=1}^2 \sum_{j=1}^3 \rho_{hj} \ln y_{h,i,t} \ln w_{j,i,t} + \sum_{h=1}^2 \omega_h \ln y_{h,i,t} \ln E_{i,t} \\
 & + \sum_{h=1}^2 \varphi_h T \ln y_{h,i,t} + \sum_{j=1}^3 \theta_j \ln w_{j,i,t} \ln E_{i,t} \\
 & + \sum_{j=1}^3 \vartheta_j T \ln w_{j,i,t} + \tau_1 T \ln E_{i,t} + \pi_0 GFC + \sum_{h=1}^2 \pi_h GFC \ln y_{h,i,t} \\
 & + \pi_3 GFCt + \sum_{j=1}^3 \vartheta_j GFC \ln w_{j,i,t} \\
 & + \chi_0 SDC + \sum_{h=1}^2 \chi_h SDC \ln y_{h,i,t} + \chi_3 SDCt + \sum_{j=1}^3 \varepsilon_j SDC \ln w_{j,i,t} \\
 & + \psi_i + v_{it} + \eta_i + u_{it}
 \end{aligned} \tag{4}$$

where i denotes the cross-sectional unit, t denotes the time period, ϕ_c is country dummy (equal to 1 when the bank is located in country c , and 0 otherwise), $y_h (h = 1, 2)$ is output, $w_j (j = 1, 2, 3)$ are input prices, $\ln E_t$ is the natural logarithm of the capital ratio, GFC and SDC are dummy variables that take the value 1 for the periods 2007–2008 and 2010–2012, respectively, and 0 elsewhere, and t is a time trend.⁹

To guarantee linear homogeneity in factor prices, we assume the following:

$$\sum_{j=1}^3 \beta_j = 1; \sum_{j=1}^3 \gamma_{kj} = 0 \forall k; \sum_{j=1}^3 \rho_{hj} = 0 \forall h \tag{5}$$

to implement linear homogeneity into the trans-log cost function, it is necessary and sufficient to apply the following standard symmetry restrictions:

$$\delta_{hk} = \delta_{kh} \forall h, k \text{ and } \gamma_{jk} = \gamma_{kj} \forall j, k \tag{6}$$

⁸ Maudos et al. (2002), Lensink et al. (2008) and Lozano-Vivas and Pasiouras (2010) did not include a trend in the cost function. This would assume that the frontier is constant over time and consequently all the productivity changes would be attributed to changes in cost efficiency or changes in economies of scale.

⁹ Hughes and Mester (2013) consider equity capital as a production input and argue that omitting it from the cost function leads to a mis specified cost function.

therefore, to impose linear homogeneity restrictions, we normalize the dependent variable and all input prices by the price of labour (w_l).

When estimating inefficiency for a large group of euro area banks, the question arises whether to estimate a common frontier for all banks or rather country-specific frontiers. The latter is usually justified when country specific circumstances affect the best practise banks. However, estimating country-specific frontiers is challenging for some euro area countries where there are not enough data for a meaningful estimation using the parametric approach. Also, integration and liberalisation of banking services in the context of the single monetary area, the single passport for financial services and recent progress with the European banking union speak in favour of estimating a single frontier, notwithstanding the fact that the operating environment for banks in the euro area remains somewhat heterogeneous.¹⁰ Bos and Schmiedel (2007) test whether commercial banks in 15 European countries share a common cost and profit frontier over the period 1993–2004. Their findings indicate a common technology, which is supportive of a single and integrated European banking market. Also, Goddard et al. (2013) find that the integration of EU financial markets following the introduction of the euro in 1999 and the harmonisation in european financial regulation were instrumental in increasing the intensity of bank competition.

A related question is whether the frontier should be estimated for different types of banks (commercial banks, saving banks, cooperative banks, etc.). A global frontier allows comparison of efficiency of different ownerships relative to the best practice in the sector, whereas the latter only permits comparison of efficiency among the same ownership. Hence, we follow Altunbas et al. (2007) and estimate a global cost-efficiency frontier for all ownerships and countries in the sample.

5 Data

Our dataset consists of a panel of commercial, cooperative and savings banks for the period 2006–2017 gathered from BankFocus.¹¹ We follow Casu et al. (2016) and include a bank in our sample until it is absorbed by another one. As such, they are treated as separate units prior to the merger. Banks are classified as commercial if they are mainly active in retail, wholesale and private banking (*i.e.*, universal banks). Savings and cooperative banks are mainly active in retail banking (with the latter having a cooperative

¹⁰ See Fiordelisi and Molyneux (2006) for a discussion on common versus country-specific frontier analysis.

¹¹ Data were collected via BankFocus (previously Bankscope) based on Moody's (previously Fitch).

Table 1 Minimum and maximum number of banks per country and business model during the period 2006–2017

Countries	Business model			All banks
	Commercial banks	Cooperative banks	Savings banks	
Austria	13/20	32/56	43/74	88/150
Belgium	6/11	2/4	1/3	9/18
Cyprus	1/9	1/2	1/1	3/12
Estonia	1/1	0/0	0/0	1/1
Finland	2/13	1/3	1/8	4/24
France	41/51	39/61	5/11	85/123
Germany	23/44	555/656	343/425	921/1125
Greece	2/5	1/1	0/0	3/6
Ireland	1/5	0/0	0/0	1/5
Italy	30/38	237/314	13/22	280/374
Luxembourg	3/13	1/7	1/1	5/21
Malta	2/5	0/0	0/0	2/5
Netherlands	3/13	1/1	1/1	5/15
Portugal	2/7	1/4	1/74	4/85
Slovakia	3/7	0/0	1/2	4/9
Slovenia	5/9	2/2	1/1	8/12
Spain	7/19	8/47	3/11	18/77
Total EA	145/270	881/1158	415/634	1441/2062

Source: Author's calculations based on BankFocus

ownership structure).¹² In order to drop institutions with unreliable or low-quality data or banks that might have been misclassified, we removed banks that: *i*) Recorded a change in the gross value of total assets of more than 50% in a particular year; *ii*) Reported negative loans or securities; *c*) Reported deposits higher than total assets; *iii*) Reported total costs (without value adjustments) above 30% of assets; *iv*) Have a gross loans-to-total assets ratio below 33% or above 90% -to remove institutions that do not provide loans to the economy or that serve as SPVs-; and *v*) Hold average assets for the whole period of below EUR 50 million (small banks). After applying these rules, our sample consists of an unbalanced panel of between 1441 and 2062 banks (depending on the year) from 17 euro area countries.

The distribution of banks by business model and country is presented in Table 1. The Table shows that more than half of the banks in our sample are located in Germany. The reason is that Germany has a large system of cooperative and savings banks. Other countries with a relatively large presence in the sample are Italy (large number of cooperative banks) and Austria (savings banks). Regarding business models, the Table shows that most banks are cooperative and savings banks.

Table 2 presents key features of banks by business model. As expected, commercial banks are, on average, the largest

institutions (holding average assets of EUR 68.9 billion at end-2017). They also possess (on average) the largest share of loans to total assets (approximately 66.0%), while the share of other earning assets is broadly comparable across banks. Cooperative and savings banks are relatively more dependent on customer deposits, while commercial banks have a somewhat more diversified source of funding. Commercial banks also seem to recruit more expensive – and probably more skilled – staff, as they tend to offer a wider range of products to a broader range of customers, often also in foreign countries. Also, commercial banks' average costs are higher. Differences in equity to assets, the price of physical capital and the price of funds, are relatively small. However, differences among banks are significant, even within the same group of banks.

There is a long-standing discussion in the literature regarding the distinction between bank outputs and inputs, particularly about the classification of deposits as inputs or outputs. The *production approach* considers that banks provide services related to both loans and deposits. The idea is that deposits are attractive because they generate fee and commission income, they generate relationships with clients, and they are a stable and cheap source of funding. As such, banks devote resources to the origination and management of deposits.¹³ By contrast, according to the *intermediation approach* banks use liabilities (deposits, equity, debt) to produce assets (loans, securities, other yielding assets, etc.). A drawback of the former is that it

¹² Other business models, such as real estate and mortgage banks were not included in the sample despite the importance of real estate financing in the euro area. Given that these banks are also involved in project development, their financial ratios are difficult to compare with the three categories considered in this analysis.

¹³ See Berger et al. (1987) and Camanho and Dyson (2005) for a discussion.

Table 2 Key features of banks per business model (*end-2017*)

			Total assets (bn. of euros)	Loans to assets	Other earning assets to assets	Customer deposits to assets	Price of labour (th. of euros)	Price of funds	Price of physical capital	Equity to assets	Average cost
Business model	Commercial	Min.	0.1	34.2%	0.0%	0.8%	44.1	0.1%	27.2%	5.7%	0.7%
		Mean	68.9	65.7%	25.8%	71.5%	75.2	0.6%	114.8%	9.6%	3.0%
		St.dev.	220.0	13.6%	12.6%	22.5%	17.8	0.4%	62.8%	3.5%	2.0%
		Max.	1960.0	89.6%	66.0%	99.0%	95.4	1.2%	223.4%	15.7%	17.2%
	Cooperative	Min.	0.1	33.0%	4.6%	14.8%	44.1	0.1%	27.2%	5.7%	0.8%
		Mean	5.6	61.7%	35.4%	78.5%	64.1	0.5%	81.5%	9.9%	2.3%
		St.dev.	63.3	12.2%	12.5%	15.2%	11.6	0.3%	47.0%	2.4%	0.9%
	Savings	Min.	0.1	33.1%	5.6%	27.1%	44.1	0.1%	27.2%	5.7%	1.2%
		Mean	6.7	63.3%	32.9%	84.2%	58.6	0.5%	113.5%	9.9%	2.6%
		St.dev.	89.0	12.3%	12.7%	9.7%	9.7	0.3%	57.9%	2.3%	0.6%
	All euro area	Min.	0.1	33.0%	0.0%	0.8%	44.1	0.1%	27.2%	5.7%	0.7%
		Mean	14.4	62.8%	33.3%	79.5%	63.9	0.5%	95.1%	9.9%	2.5%
St.dev.		108.0	12.5%	13.0%	15.4%	13.1	0.3%	54.7%	2.5%	1.0%	
		Max.	2130.0	89.6%	68.3%	99.4%	95.4	1.2%	223.4%	15.7%	22.0%

Notes: The price of labour is calculated as personnel expenses over the total number of employees; the price of physical capital is calculated as the ratio of other overhead costs to non-earning assets; and the price of funds is computed as the ration between interest costs and total liabilities

Source: Author's calculations based on BankFocus

restricts the definition of total costs (because deposits are not considered an input) and limits the measurement of efficiency to operational costs. Consequently, under the reasonable assumption that banks invest resources in optimising their funding structure, the efficiency results would be distorted. Moreover, the savings banks in our sample provide products, and not only services and some are even larger than commercial banks. Hence, we adopt the *intermediation approach* of Sealey and Lindley (1977) and treat liabilities as inputs and assets as outputs. In particular, we view banks as firms that use labour, other aggregate inputs and financial capital to produce loans and other earning assets.¹⁴

Regarding the price of inputs, we compute the price of labour as labour expenses over the number of employees.¹⁵ For the price of other aggregate inputs, we use the ratio of other (non-labour) administrative costs to fixed assets. The price of funds is computed as the ratio between interest expenses and total liabilities. Total costs, our dependent variable, is computed

as the sum of these three components. By including interest costs (cost of financing) we capture a more comprehensive overview of banks' business profiles.¹⁶ This specification of outputs and inputs is similar to most of the previous studies. Most of the literature has estimated cost functions with the same inputs while the number of outputs has varied from two to five.¹⁷ Finally, we follow Berger and Mester (1997), Hughes and Mester (2008), Fiordelisi et al. (2011), Boucinha et al. (2013) and Duygun et al. (2015) and employ equity to total assets as a quasi-fixed input to control for differences in risk preferences. Table 3 presents descriptive statistics for the variables included in the model.

6 Econometric results

The econometric results of the estimation of Eq. (4) subject to restrictions (5) and (6) are reported in Table 4. As

¹⁴ Boucinha et al. (2013) test for the inclusion of deposits as outputs following the methodology developed by Hughes and Mester (1993). Implementation of such test requires a breakdown of interest costs into those paid on deposits versus on other liabilities. Such granular data are not available for the sample under consideration.

¹⁵ Part of the literature computes the price of labour as the ratio between personell expenses and total assets. By calculating the price of labour relative to total assets one would actually capture labour productivity as well (Maudos et al., 2002).

¹⁶ Altunbas et al. (2007) also compute total costs including operating and financial costs.

¹⁷ A few studies that have estimated a cost function with the same inputs are Altunbas et al. (1999), Altunbas et al. (2001), Maudos et al. (2002), Altunbas et al. (2007), Feng and Serleis (2009), Fiordelisi et al. (2011), Boucinha et al. (2013) and Tsionas and Kumbhakar (2014). Altunbas et al. (2001) focus on five outputs, namely mortgage loans, public loans, other loans, aggregate securities and off-balance sheet items.

Table 3 Variables included in the cost function

	Unit	Obs.	Mean	Std. Dev.	Min	Max
Dependent variable						
Total costs	Mill. EUR	21,224	352	2827	−1	101000
Outputs						
Gross loans	Mill. EUR	21,224	6384	47,200	9.117363	1,220,000
Other earning assets	Mill. EUR	21,224	4325	41,200	0.009	1,180,000
Input prices						
Personnel costs per employee	000 EUR	19,332	58.4	12.0	38.7	95.4
Interest expenses to total liabilities	%	21,224	2%	1%	0%	4%
Other overheads to non-earning assets	%	19,829	85%	44%	27%	225%
Semi-fixed input						
Total equity to total assets	%	21,224	9%	3%	4%	16%

Source: Author's calculations based on BankFocus data

mentioned before, we normalize the dependent variable (total costs) and all input prices by the price of labour (w_l). Four models have been estimated. The first model (M1) is the standard trans-log cost function with three inputs and two outputs, a linear and a quadratic time trend, and the bank capital ratio. To control for unobserved heterogeneity across countries, the second model (M2) includes also country dummies (equal to 1 for all the banks located in the same country and 0 otherwise).

The third model (M3) extends Model 1 by including dummies for the global financial crisis (GFC) and the sovereign debt crisis (SDC) and the respective interaction terms. The dummy for the GFC and the SDC is 1 for 2007–2008 and 2010–2012, respectively, and 0 otherwise. Other studies have also included dummies to account for changes in regulation, the financial crisis and the euro introduction (Badunenko and Kumbhakar 2017; Casu et al. 2016). Finally, the fourth model (M4) includes country dummies and dummies for the GFC and the SDC. As such, M4 is the most encompassing model.

Input and output point elasticities have the expected signs in all the models. Outputs have positive coefficients and are significant at the 1% level, implying that generating more output is costly. The coefficient of loans (the first output) ranges from 0.07 to 0.15% and that one securities ranges between 0.6 and 0.67%, depending on the model. Moreover, the outputs squared are also positive and significant at the 1% level, implying that costs increase exponentially with output. Also the product between the two inputs is negative, suggesting that more diversified banks tend to have lower costs. Between the two relative input prices included in the equation, only the coefficient for the price of funds (w_2/w_l) is positive and significant (at the 1% level).

The capital ratio has a negative impact on costs (significant at the 1% level), which captures the fact that better capitalised banks pay a smaller risk premium (see the discussion below on the Equity-ratio effect). Interestingly, the

impact of holding higher equity ratios is non-linear, but only when controlling for the GFC and the SDC, implying that one percent increase in equity has a smaller impact on total (relative) costs when the equity ratio is already high. The coefficient for the linear trend is negative and significant at the 1% level in all models, implying that costs tend to decline at a rate of 0.04% per year. The impact of time becomes more strongly when controlling for the GFC and the SDC. The negative coefficient on the cross product of loans and securities is significant at the 1% and stable across models, suggesting the presence of economies of scope in the joint production of both outputs.

Both the GFC and the SDC are found to have increased total (relative) costs. The coefficient for the GFC stands at 0.37 and 0.36% in models M3 and M4, respectively and than one for the SDC at 0.24 in both models. Also, producing loans appears to have been more expensive during the GFC but not during the SDC (the interaction term between loans and the GFC is positive and significant but it is insignificant for the SDC). Also, the GFC appears to have exacerbated the impact of labour and funding costs compared with the SDC (the coefficients associated with the two inputs are positive and larger in the case of the GFC). Finally, the interaction of the two crises dummies with time is insignificant.¹⁸

The explanatory power of all the models is largely satisfactory (the adjusted R² is 0.987, unchanged across models). Moreover, the two crisis dummies and the respective interaction terms are significant in Model 4, both individually and jointly, suggesting that they contributed to

¹⁸ Model M4 has also been estimated including environmental variables: real GDP growth, HICP inflation, population density and percentage of population using internet banking. The results of the estimation of the efficiency scores remain largely unchanged, as the methodology deals already with heterogeneity to a large degree (by accounting for bank-specific effects and country heterogeneity). These results are available from the authors upon request.

Table 4 Econometric results for the translog cost function (*Dependent variable: $\ln(TC/w_1)$*)

	M1	M2	M3	M4
Const.	2.12***	2.055***	1.488***	1.433***
$\ln Y_1$	0.149***	0.144***	0.082**	0.074**
$\ln Y_2$	0.596***	0.602***	0.663***	0.669***
$\ln(w_2/w_1)$	0.907***	0.904***	0.767***	0.763***
$\ln(w_3/w_1)$	0.070	0.079*	0.051	0.060
$\ln Z$	-0.247***	-0.254***	-0.258***	-0.262***
t	-0.035***	-0.037***	-0.042***	-0.044***
$\ln Y_1^2$	0.124***	0.124***	0.123***	0.124***
$\ln Y_2^2$	0.118***	0.118***	0.118***	0.118***
$\ln Y_1 * \ln Y_2$	-0.221***	-0.222***	-0.219***	-0.22***
$\ln W_1 * \ln W_2$	0.084***	0.084***	0.092***	0.093***
$\ln W_1 * \ln W_3$	-0.104***	-0.11***	-0.088***	-0.093***
$\ln W_2 * \ln W_1$	-0.111***	-0.109***	-0.091***	-0.089***
$\ln W_2 * \ln W_3$	-0.056***	-0.058***	-0.056***	-0.058***
$\ln W_3 * \ln W_1$	0.009	0.008	0.005	0.003
$\ln Z^2$	0.010	0.01	0.017**	0.016**
r^2	0.002***	0.002***	0.002***	0.002***
$\ln Y_1 * \ln W_2$	-0.037***	-0.038***	-0.045***	-0.047***
$\ln Y_1 * \ln W_3$	0.000	0.001	0.001	0.001
$\ln Y_2 * \ln W_1$	-0.048***	-0.048***	-0.056***	-0.056***
$\ln Y_2 * \ln W_3$	0.002	0.002	0.002	0.002
$\ln Y_1 * Z$	0.006	0.005	0.006	0.005
$\ln Y_2 * Z$	0.012**	0.011**	0.013**	0.012**
$\ln Y_1 * t$	-0.003***	-0.003***	-0.003***	-0.004***
$\ln Y_2 * t$	0.004***	0.004***	0.004***	0.004***
$\ln W_1 * Z$	-0.138***	-0.132***	-0.101***	-0.097***
$\ln W_2 * Z$	-0.074***	-0.075***	-0.058***	-0.059***
$\ln W_3 * Z$	-0.017***	-0.018***	-0.01	-0.01*
$\ln W_1 * t$	-0.008***	-0.007***	-0.003	-0.003
$\ln W_2 * t$	-0.002*	-0.002	-0.001	-0.001
$\ln W_3 * t$	0.000	0.00	0.00	0.00
$\ln Z * t$	-0.006***	-0.006***	-0.004***	-0.004***
GFC			0.369***	0.360***
$\ln Y_1 * GFC$			0.017***	0.018***
$\ln Y_2 * GFC$			-0.012***	-0.013***
$\ln W_1 * GFC$			0.164***	0.163***
$\ln W_2 * GFC$			0.136***	0.135***
$\ln W_3 * GFC$			0.006	0.006
$GFC * t$			-0.001	-0.001
SDC			0.236***	0.237***
$\ln Y_1 * SDC$			0.003	0.004
$\ln Y_2 * SDC$			-0.008***	-0.008***
$\ln W_1 * SDC$			0.043***	0.039***
$\ln W_2 * SDC$			0.036***	0.034***
$\ln W_3 * SDC$			0.011***	0.011***

Notes: The table reports the results of the estimation of Eq. (4). The variables are defined as follows: y_1 = gross loans; y_2 = securities; w_1 = price of labour; w_2 = price of funds; w_3 = price of fixed inputs; t = time trend; z = equity-to-assets ratio; GFC = global financial crisis; SDC = sovereign debt crisis. Joint significance for country dummies is tested with a Chi2 test (p-value reported in parentheses). The LR test is a likelihood ratio test for the comparison of models M3 and M4

affect total (relative) costs in the sample under consideration. As a result, in what follows we will focus the analysis on Model 4, the most encompassing model.

7 Computing Total Factor Productivity growth (TFPG) in the euro area banking sector

This section computes Total Factor Productivity (TFP) growth in the euro area banking sector, *i.e.*, the growth in output not explained by the growth in the amount of inputs utilised. It is based on Eq. (1) presented in Section 3. As mentioned before, the estimated trans-log cost function (Eq. (4) above) can be employed to compute each of the components of Total Factor Productivity Growth, namely overall technical efficiency, technological progress and the equity and the scale effect.

The first Sub-section (7.1) looks at technical efficiency in the euro area banking sector. Sub-section 7.2 presents the results of the estimation of technological progress. The Sub-section presents the temporal evolution of technological progress and its three components, namely pure, scale-augmenting and non-neutral technological progress. Sub-section 7.3 presents the third component of Total Factor Productivity Growth (TFPG) and is based on the estimation of the shadow cost of equity. The scale effect is presented in Sub-section 7.4 and requires the previous calculation of economies of scale and the shadow cost of equity. Finally, the four components are taken together to compute Total Factor Productivity Growth in Subsection 7.5.

7.1 Technical efficiency

The overall technical efficiency of the euro area banking sector, together with the estimation of persistent and time-varying efficiency across bank's business models is reported in Table 5. On average for all banks, the persistent component of efficiency amounted to about 88.1%, while the time-varying efficiency component amounted to about 95.5% during the period from 2006 to 2017.¹⁹ These findings suggest that the median bank uses 11.9 and 4.5% more resources than the bank that is at the efficiency frontier, due to persistent and time-varying factors, respectively. Hence, after controlling for bank heterogeneity and persistent efficiency, the share attributed to time-varying efficiency is relatively small. Put otherwise, structural long-term factors (such as location, client structure, macroeconomic environment, regulation, etc.) seem to play a bigger role for bank efficiency than factors that change over time.

There seems to be little difference in time-varying efficiency across business models, while the differences are larger for persistent efficiency. In particular, cooperative and saving banks are found to be more efficient than

¹⁹ Changes in persistent efficiency over time reflect a change in the underlying market shares of the banks in the sample and changes in the sample per se.

Table 5 Efficiency per bank specialization (*median for all banks and each category*)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persistent efficiency												
Commercial	84.4%	84.5%	84.4%	84.2%	84.4%	84.5%	85.2%	85.3%	85.2%	84.9%	85.3%	85.3%
Cooperative	89.4%	89.4%	89.4%	89.4%	89.4%	89.4%	89.3%	89.4%	89.4%	89.0%	89.4%	89.4%
Savings	85.8%	85.8%	86.0%	86.0%	86.2%	86.3%	86.5%	86.4%	86.4%	86.1%	86.3%	86.2%
All banks	88.2%	88.1%	88.2%	88.2%	88.2%	88.2%	88.2%	88.1%	88.2%	87.8%	88.1%	88.0%
Time-varying efficiency												
Commercial	96.2%	96.2%	95.9%	96.0%	95.7%	95.4%	95.0%	95.4%	95.4%	95.6%	94.9%	94.7%
Cooperative	95.3%	95.4%	95.3%	95.3%	95.3%	95.3%	95.4%	95.5%	95.6%	95.4%	95.8%	96.4%
Savings	95.6%	95.2%	95.7%	95.5%	95.4%	95.7%	95.7%	95.7%	95.7%	95.7%	94.3%	95.4%
All banks	95.5%	95.4%	95.4%	95.4%	95.4%	95.5%	95.5%	95.6%	95.6%	95.5%	95.2%	95.9%
Overall efficiency												
Commercial	79.9%	81.4%	80.7%	80.5%	81.1%	80.2%	80.8%	81.0%	80.9%	81.2%	80.1%	79.8%
Cooperative	84.8%	84.9%	85.0%	84.9%	84.9%	84.9%	85.1%	85.1%	85.1%	84.5%	85.1%	85.7%
Savings	81.8%	81.6%	82.2%	81.9%	82.1%	82.6%	82.8%	82.7%	82.5%	82.1%	80.6%	81.5%
All banks	83.6%	83.7%	83.8%	83.9%	84.0%	84.1%	84.1%	84.1%	84.1%	83.6%	83.4%	83.9%

Note: The relative distance to the frontier for persistent and time-varying inefficiency is computed based on v_{it}^+ and η_{it}^+ , respectively (as described in Eq. 4)

Source: Author's calculations based on BankFocus data

commercial banks. This finding suggests that differences across banks' business models play a more important role for efficiency than year-on-year changes in management decisions. This result is somewhat expected, because commercial banks are larger, and as such, they incur larger sunk costs. This evidence is supported by Altunbas et al. (2001), who find that mutual and saving banks in Germany have cost efficiency advantages over their private commercial banking counterparts.

Overall bank efficiency, computed as the product between persistent and time-varying efficiency, was around 84% for the entire euro area banking sector over the period from 2006 to 2017. These findings are in line with those for US commercial banks (Feng and Serletis 2009), Portuguese banks (Boucinha et al. 2013), German banks (Altunbas et al. 2001) and a sample of European banks (Maudos et al. 2002). Our findings suggest slightly higher cost efficiency than the studies mentioned in Bos and Schmiedel (2007) and Fiordelisi et al. (2011) for European commercial banks over the period 1995–2007. In other words, our findings suggest that if the median bank would operate on the efficiency frontier, it could produce the same level of output with only around 84% of current costs.

Looking at the evolution of efficiency across bank size, as measured by the respective market share in the country of origin, larger institutions tend to display lower overall efficiency scores (Table 6). At the end of the sample (in 2017), the overall efficiency score for banks above the 75th percentile (*i.e.*, the largest banks) was around 5.9 percentage points lower than for those

below the 25th percentile.²⁰ This difference, which after a period of convergence, widened again more recently, seems to be mainly the result of differences in persistent efficiency. One reason that could explain why larger institutions are less efficient is that they are more difficult to manage, as they deploy a more sophisticated business model. On the other hand, larger institutions might invest more in other aspects of the business, such as brand value or strategy, which are focused mainly on increasing market power rather than on cost efficiency. As these activities are recorded in total costs but not in bank outputs (they are intangible assets), they might lead to lower efficiency in these institutions. By contrast, time-varying efficiency seems to be less related to size, suggesting that the lower efficiency of larger institutions is a structural, rather than a time-varying phenomenon.

Finally, we look at banks with different efficiency scores across several indicators (average cost, cost to income, etc.) (Table 7). Banks are divided into three groups: *i*) Banks with overall efficiency score below the 10th percentile (most inefficient banks); *ii*) The total sample; and *iii*) Banks with overall efficiency scores above the 90th percentile (most efficient banks). As expected, more efficient banks, according to the stochastic frontier analysis, also tend to record a lower cost-to-income ratio, higher profitability (measured by return on assets, ROA) and a larger share of high yield items in the balance

²⁰ Feng and Serletis (2009) also find that the largest commercial banks are less efficient than their smaller counterparts in a sample of US banks.

Table 6 Efficiency by bank market share in the local market (*median for each category*)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Persistent efficiency												
<25th pctile.	92.0%	91.7%	91.3%	91.4%	90.6%	90.6%	90.5%	90.9%	91.0%	90.6%	90.9%	91.0%
25th to 50th pctile.	89.3%	89.1%	88.4%	88.3%	87.4%	87.3%	87.5%	88.1%	88.2%	88.2%	88.1%	88.2%
50th to 75th pctile.	87.1%	86.7%	86.2%	86.2%	86.2%	86.2%	86.1%	86.2%	86.3%	86.2%	86.4%	86.6%
>75thpctile.	85.2%	85.4%	86.0%	86.2%	86.6%	86.5%	86.9%	86.2%	86.3%	86.0%	86.6%	85.8%
Time-varying efficiency												
<25th pctile.	96.0%	95.7%	95.5%	95.4%	95.6%	95.5%	95.4%	95.2%	95.2%	94.9%	95.6%	96.1%
25th to 50th pctile.	95.4%	95.3%	95.3%	95.4%	95.3%	95.4%	95.6%	95.5%	95.6%	95.5%	95.6%	96.2%
50th to 75th pctile.	95.2%	95.2%	95.4%	95.4%	95.3%	95.5%	95.6%	95.8%	95.8%	95.5%	95.0%	95.8%
>75thpctile.	95.5%	95.3%	95.6%	95.3%	95.2%	95.5%	95.3%	95.6%	95.8%	95.9%	94.9%	95.7%
Overall efficiency												
<25th pctile.	88.1%	87.7%	87.1%	87.0%	86.4%	86.5%	86.3%	86.3%	86.4%	85.8%	86.5%	87.1%
25th to 50th pctile.	84.8%	84.5%	84.0%	84.3%	83.1%	83.1%	83.6%	84.2%	84.2%	84.0%	83.7%	84.1%
50th to 75th pctile.	82.5%	82.2%	82.2%	81.9%	82.0%	82.2%	82.1%	82.5%	82.7%	82.3%	81.2%	82.8%
>75thpctile.	80.9%	81.1%	81.9%	82.0%	82.1%	82.6%	82.4%	82.3%	82.6%	82.4%	81.1%	81.2%

Note: The Table reports median efficiency scores (relative distance to the frontier) by bank size. The bank size is measured by the market share in the country of origin

Source: Author's calculations based on BankFocus data

sheet. Moreover, the most efficient banks tend to have lower credit risk ratios than the less efficient counterparts, which suggests that they are more efficient at handling credit risk. These results are broadly in line with those of Boucinha et al. (2013) for Portugal and Nițoi and Spulbar (2015) for Central and Eastern Europe. Lastly, more efficient banks tend to be better capitalised.²¹ On the one hand, this result suggests that higher profitability associated with higher efficiency allows banks to accumulate retained earnings and increase capital ratios. At the same time, this finding might also suggest that banks with a higher share of equity have the incentive to increase efficiency, since equity is the most expensive source of financing.²²

7.2 Technological progress

Including²³ a time trend in the cost function (linear, squared and interacted with other exogenous variables) allows for the estimation of technological progress, defined as the effect of

²¹ In this regard, it is important to note that since we included equity as a quasi-fixed input in the equation, our efficiency measures are not affected by the fact that some banks have higher equity levels and therefore a lower cost of liabilities.

²² Fiordelisi et al. (2011) find that the link between bank efficiency and solvency runs both ways: more efficient banks become better capitalized and higher capital levels tend to have a positive effect on efficiency levels in European commercial banks. Altunbas et al. (2007) and Nițoi and Spulbar (2015) find that banks with lower solvency rates are more inefficient in Central and Eastern Europe and in European cooperative banks, respectively.

²³ Results reported in the following sections are based on the estimation of a trans-log cost function estimated with SFA methods. Hence, unlike before where we computed two elements of inefficiency (persistent and time-varying), other elements of total factor productivity growth are calculated directly from the same cost function.

time on total costs and computed as the partial derivative of total costs with respect to time ($T\text{PROG} = \partial \ln TC_t / \partial t$). According to Baltagi and Griffin (1988) and Kumbhakar and Heshmanti (1996), technological progress can be divided into three components. The first is called “pure technological progress” and depends only on the time trend. The second is called “scale-augmenting technological progress” and depends on the interaction terms between time and outputs, capturing the change in the sensitivity of total costs with respect to time as output changes. The third component is called “non-neutral technological progress” and reflects the changes in the sensitivity of total costs to time, as input prices change.

The three components are presented in Fig. 1. On average between 2006 and 2017, the annual rate of technological progress for the median euro area bank amounted to 3.1%. Altunbas et al. (2007) estimated a much higher rate of technological progress for German banks over the period 1989 to 1996 (10% on average). This finding might be affected by the fact that this period overlaps by and large with the introduction of internet and new computer technologies.²⁴ Our results are also broadly in line with Altunbas et al. (1999) and Boucinha et al. (2013). The former find the rate of technological progress for banks in fifteen European countries to be between 2.8 and 3.6% over the period 1989 to 1996, while the latter estimated the rate of technological progress of Portuguese banks between 2000 and 2006 to be around 2 to 3%.

The largest component of technological progress is non-neutral, followed by pure technological progress (amounting to around 3.9 and 2.8%, respectively, on average in the whole sample for the median euro area bank). The finding

²⁴ It may also reflect a miss-specified model, as their econometric specification omits risk and cost of equity considerations.

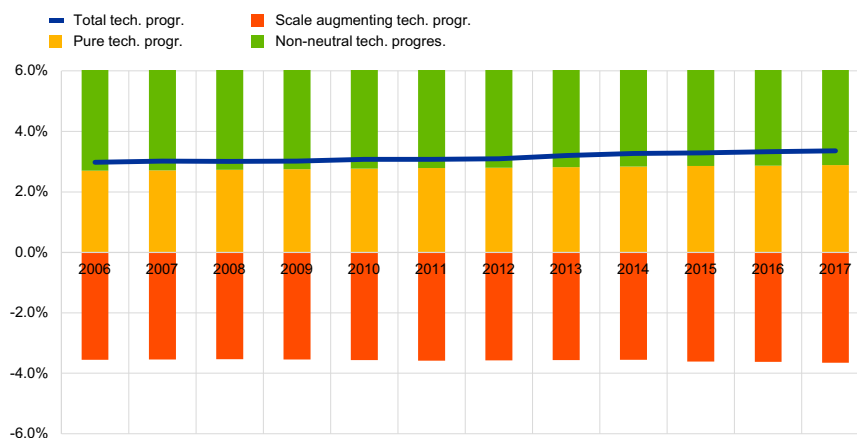
Table 7 Bank’s characteristics by overall efficiency scores (% if not stated otherwise; median)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Banks with efficiency below 10th percentile												
Average cost	4.85%	4.66%	4.89%	5.08%	4.35%	3.91%	3.95%	3.67%	3.44%	3.35%	3.26%	2.92%
Cost to income	71.43%	73.54%	73.09%	71.21%	74.31%	73.38%	72.57%	73.43%	72.94%	74.23%	84.74%	79.34%
Loans to assets	57.77%	58.93%	58.45%	58.14%	58.31%	60.64%	61.72%	59.20%	58.54%	57.23%	59.11%	59.70%
Return on assets (ROA)	0.30%	0.29%	0.26%	0.25%	0.24%	0.20%	0.19%	0.15%	0.19%	0.16%	0.14%	0.19%
Market share	0.05%	0.05%	0.04%	0.04%	0.02%	0.02%	0.02%	0.03%	0.03%	0.04%	0.06%	0.06%
Credit risk	0.78%	0.64%	0.60%	0.78%	0.60%	0.42%	0.49%	0.51%	0.40%	0.40%	0.13%	0.19%
Equity to assets	6.65%	6.59%	6.30%	6.38%	6.64%	6.91%	7.24%	7.50%	7.39%	7.94%	8.66%	8.46%
All banks												
Average cost	4.33%	4.28%	4.52%	4.78%	4.02%	3.54%	3.50%	3.21%	2.99%	2.77%	2.59%	2.38%
Cost to income	65.69%	69.29%	70.27%	68.33%	66.67%	66.39%	66.90%	66.80%	66.85%	68.29%	72.62%	71.84%
Loans to assets	63.35%	62.72%	62.08%	62.08%	62.78%	63.58%	63.26%	63.16%	62.56%	63.30%	63.19%	63.61%
Return on assets (ROA)	0.41%	0.31%	0.24%	0.28%	0.28%	0.26%	0.26%	0.24%	0.25%	0.23%	0.22%	0.23%
Market share	0.03%	0.02%	0.02%	0.02%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%	0.03%
Credit risk	0.73%	0.57%	0.63%	0.65%	0.57%	0.35%	0.31%	0.29%	0.22%	0.24%	0.12%	0.11%
Equity to assets	6.86%	6.95%	6.78%	7.01%	7.29%	7.79%	8.13%	8.56%	8.89%	9.02%	9.30%	9.55%
Banks with efficiency above 90th percentile												
Average cost	3.87%	3.94%	4.32%	4.63%	3.71%	3.16%	3.19%	2.87%	2.61%	2.48%	2.30%	2.13%
Cost to income	62.78%	65.08%	68.56%	66.67%	65.75%	64.71%	65.58%	64.47%	64.71%	65.71%	69.90%	69.16%
Loans to assets	64.35%	62.94%	63.10%	63.57%	64.98%	65.35%	64.12%	64.32%	63.93%	66.18%	65.89%	65.55%
Return on assets (ROA)	0.52%	0.49%	0.28%	0.31%	0.36%	0.34%	0.33%	0.33%	0.35%	0.36%	0.32%	0.33%
Market share	0.03%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.01%	0.03%	0.03%	0.03%	0.04%
Credit risk	0.52%	0.47%	0.57%	0.60%	0.54%	0.48%	0.41%	0.26%	0.24%	0.22%	0.09%	0.06%
Equity to assets	7.39%	7.86%	7.20%	7.46%	7.80%	8.26%	8.65%	8.99%	9.20%	9.31%	9.51%	9.87%

Notes: The average cost is the ratio between total costs and total assets. Cost to income is the ratio between operating costs and operating income. Credit risk is the ratio between loan loss provisions and total assets. The Z-score is calculated as: $Z = (EAR + ROA^{3ya}) / \sigma_{ROA}^{3ya}$, where EAR is the equity-to-asset ratio for the current period, ROA^{3ya} is the 3-year moving average of ROA and σ_{ROA}^{3ya} is the 3-year standard deviation of ROA

Source: Author’s calculations based on BankFocus data

Fig. 1 The rate of technological progress and its components for the euro area banking sector (median). Note: Technological progress leads to a decline in total costs. Source: Author’s calculations based on BankFocus data

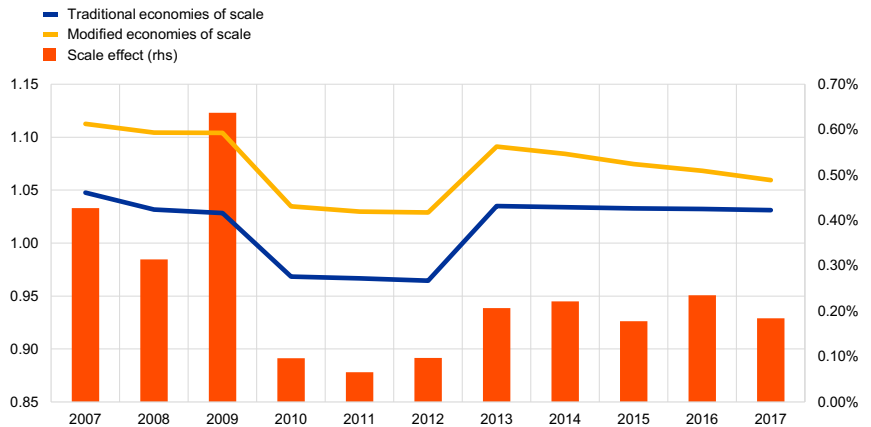


for pure technological progress suggests that costs tend to decrease over time, holding constant the efficient scale of production for our two outputs and the shares of each input in total costs. By contrast, the scale-augmenting component, or the sensitivity of total costs with respect to variations in

the efficient scale of production, contributed to increase costs over the period (average value of about −3.6%).

The rate of technological progress increased in our sample, from 3.0% in 2006 to 3.4% in 2017. The main drivers were an increase in non-neutral and pure technological progress (by 0.3

Fig. 2 Shadow cost of the equity ratio and the cost of equity derived from the CAPM (*weighted average*). Source: Author's calculations based on BankFocus data



and 0.2 p.p., respectively), which more than compensated the decline in scale augmenting technological progress (by 0.1 p. p.). Decreasing costs of funding in the euro area banking sector are behind the increase in the rate of change of non-neutral technological progress. This observation implies that technological progress is mostly driven by factors that are outside the control of the banks. Importantly, technological progress does not seem to exhibit a different pattern in the two crises compared with other periods in the sample.²⁵

7.3 The equity-ratio effect

The equity effect captures the impact of the cost of equity and changes in equity (the product of the two) on Total Factor Productivity Growth. It measures the impact of a change in equity on bank costs in a particular year, weighted by the change in equity.

Including equity capital in the cost function is important to avoid biased efficiency estimates, since: (i) Equity is a source of funding and should be considered a specific, quasi-fixed input; (ii) The new regulatory regime since Basel III requires higher capital requirements, influencing the production and cost profile of banks; and (iii) Holding more equity could lead to lower total costs (Dijkstra 2013; Hughes and Mester 2013), as the price of equity falls when the amount of equity increases (Modigliani and Miller 1958). Moreover, better capitalised institutions will be perceived as less risky and creditors could reward them by charging less interest on other liabilities. Also, higher capital levels can act as a signalling device, resulting in banks paying lower prices for other inputs. Therefore, this cost reduction should not be confused with technical efficiency (Hughes et al. 2001; Hughes and Mester 2013).²⁶

As such, Hughes and Mester (2013) suggest including the cost of equity capital (*i.e.*, the return on capital times the capital

held by the bank) in the cost function. The authors note that when a bank is publicly traded, the return on capital can be computed from an asset pricing model and included in the cost function. However, when most banks are not publicly traded (as in our case), a cost function such as Eq. (4) including equity capital can be employed to obtain a shadow price of equity capital. We follow Berger and Mester (1997), Bos and Schmiedel (2007), Hughes and Mester (2008), Shen et al. (2009), Fiordelisi et al. (2011), Boucinha et al. (2013) and Duygun et al. (2015) and include the equity ratio in the cost equation. The shadow cost of the equity ratio (SCOER) can then be computed as the negative of the partial derivative of the cost function with respect to the equity ratio and shows the cost savings associated with an increase in the equity ratio.²⁷

For comparison, Fig. 2 presents our SCOER and the cost of equity derived from a Capital Asset Pricing Model (Markowitz 1952) available from data providers.²⁸ Our shadow cost of equity is comparable with results from the CAPM and points to a sharp increase after the start of the global financial crisis, peaking in 2009 at about 7.6% and decreasing afterwards, to 2.7% in 2017. These results suggest that the reward for being better capitalised increased in times of financial stress, but decreased afterwards, as equity ratios increased for euro area banks. The results for the overall level of the shadow cost of equity are in line with those found in the literature. For example, Shen et al. (2009) estimate a shadow cost of equity ranging from 2 to 6% for a sample of Asian countries, while Boucinha et al. (2013) estimate it to range from 1 to 20% for Portuguese banks. Other studies that included equity in the cost function did not compute a shadow cost of equity (Maudos et al. 2002; Bos and Schmiedel 2007; Koetter and Poghosyan 2009; Fiordelisi et al. 2011; Casu et al. 2016).

²⁵ This is seen by the fact that the interaction between the crises dummies and the time trend is not significant.

²⁶ Altunbas et al. (1999 and 2007) and Altunbas et al. (2001) estimate a trans-log cost function for European and German banks, respectively, but omit equity from the estimated equation.

²⁷ Duygun et al. (2015) compute the SCOER to estimate its contribution to total factor productivity change in emerging market economies and call it the *regulatory constraint effect*.

²⁸ Boucinha et al. (2013) compare their SCOER with short-term money market interest rates and the implicit price of funding.

Table 8 Modified returns to scale, by size and bank specialisation (*median by group*)

	Size				Specialisation		
	Size 1	Size 2	Size 3	Size 4	Commercial	Cooperative	Savings
2006	1.13	1.12	1.11	1.10	1.10	1.12	1.11
2007	1.13	1.12	1.11	1.10	1.10	1.11	1.11
2008	1.11	1.11	1.10	1.09	1.09	1.11	1.10
2009	1.11	1.11	1.10	1.09	1.09	1.11	1.10
2010	1.04	1.03	1.03	1.02	1.02	1.04	1.03
2011	1.04	1.03	1.02	1.01	1.01	1.03	1.03
2012	1.04	1.03	1.02	1.01	1.01	1.03	1.03
2013	1.10	1.09	1.09	1.07	1.08	1.09	1.09
2014	1.09	1.09	1.08	1.07	1.07	1.08	1.08
2015	1.09	1.08	1.07	1.06	1.06	1.08	1.07
2016	1.08	1.07	1.07	1.06	1.05	1.07	1.07
2017	1.07	1.06	1.06	1.05	1.05	1.06	1.06

Note: The bank size is measured by the respective market share in the country of origin. Commercial, cooperative and saving banks are defined in Section 4

Source: Author's calculations based on BankFocus data

7.4 Scale effect

The fourth component of Total Factor Productivity Growth (TFPG) is the so-called “scale effect”. It is computed as the product between returns to scale minus one and (weighted) output growth. This component captures the importance of operating at the optimal scale (Kumbhakar et al. 2014 and 2015). Economies of scale are typically computed as the inverse of the output cost elasticity based on the trans-log cost function.²⁹ For the trans-log cost function that we use in this analysis, the output cost elasticity is observation-specific (*i.e.*, it varies by bank and over time).

When calculating economies of scale in this traditional manner, the implicit assumption is a constant cost of equity. However, Hughes and Mester (2013) find that ignoring equity capital from the calculation of scale economies leads to erroneous measures of scale economies. As a result, a modified measure of the scale elasticity that accounts for the impact of the cost of equity can be computed, as follows:³⁰

$$SEL_{EQ} = SEL / (1 - SCOE) \quad (7)$$

where SEL_{EQ} is the modified measure of the scale elasticity, SEL is the traditional measure of scale elasticity and $SCOE$ is the shadow cost of the equity ratio, computed in the previous section.

The modified measure of economies of scale by bank size (grouped into four categories) and business model is reported in Table 8. Results show that they tend to be larger for the

smaller institutions, although the largest institutions (those in category number four) also experience economies of scale.³¹ The smallest and the largest institutions exhibited average economies of scale of about 1.09 and 1.06 over the sample, respectively. Regarding bank specialisation, economies of scale seem to be slightly larger for cooperative banks. These findings are consistent with those in Altunbas et al. (2001) for German banks, who report higher economies of scale for mutual and saving banks than for larger, commercial banks.

The modified measure of economies of scale, the constant measure of economies of scale and the scale effect are reported in Fig. 3. Our results suggest that euro area banks exhibited economies of scale of around 8%, on average over the period. By comparison, the standard measure is lower, signalling economies of scale of around 2% on average. Importantly, both measures of economies of scale were stable until 2009 and then declined during the sovereign debt crisis (SDC), before increasing again afterwards. The modified measure of economies of scale reached approximately 6% at the end of the sample. These findings suggest that increasing outputs by a factor of one in 2017 led to an increase in total costs by a factor of 0.94.³² The decline observed in the modified measure of economies of scale during the SDC is due to a decline in the shadow cost of equity presented before, and also to a decline in the traditional measure of economies of scale. In this regard, the difference between the traditional and the modified measure of economies of scale is visible during the SDC, as the former implies diseconomies of scale but not the latter.

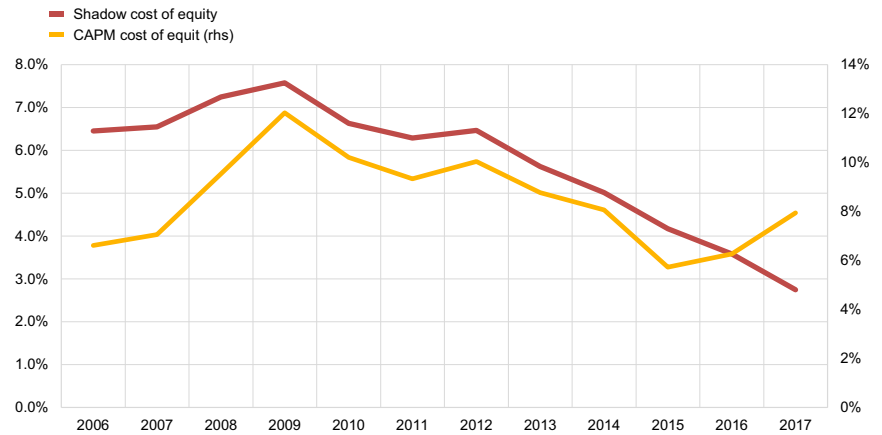
²⁹ The output cost elasticity shows the sensitivity of total costs to changes in output (*i.e.*, the sum of the partial derivatives of total costs with respect to each of the outputs: $E_{cy} = \sum_{h=1}^2 \partial \ln TC / \partial \ln y_h$).

³⁰ See Equation (14) in Hughes and Mester (2013).

³¹ The finding that smaller banks gain more from growing more is in line with Hughes and Mester (2013).

³² Altunbas et al. (2001) found higher average return to scales for a sample of German banks between 1989 and 1996, standing at about 11%, even assuming a constant cost of equity.

Fig. 3 Modified economies of scale and scale effect in euro area banks (*median*). Source: Author's calculations based on BankFocus data and Bloomberg (CAPM)



Finally, the scale effect (product between modified economies of scale and –weighted- output growth) peaked in 2009, decreased during the sovereign crisis and then increased again when bank products (loans and investments) rebounded from the sovereign debt crises. However, the scale effect did not reach again the levels observed during and before the global financial crisis. The average scale effect stood at about 0.21% of total costs after the sovereign debt crisis.

7.5 Total factor productivity growth

Having computed technical efficiency, technological progress, the equity and the scale effect, we are now able to compute the Total Factor Productivity Growth of euro area banks based on Eq. (1). Results reported in Fig. 4 show that Total Factor Productivity of the median euro area declined from above 2.6% in 2007 to below 1.7% in 2017 (*i.e.*, banks were able to generate the same amount of output with 1.7% less costs per year in 2017, compared with 2.6% in 2007). In particular, Total Factor Productivity Growth decreased after the GFC, from 2.8% to around 2.2% and remained at those levels during the SDC. Thereafter, productivity increased slightly in 2013 but then continued to decline. Casu et al. (2016) also found that the financial crisis led to a decline in productivity in euro area banks.

As traditionally found in the literature, the largest component of Total Factor Productivity (in terms of absolute size) is technological progress (Boucinha et al. 2013; Casu et al. 2016). The contribution of this component increased during the last ten years, from 3% in 2007 to 3.4% in 2017. Both the scale and equity effect also contributed to boost productivity in the banking sector (except for the latter in 2008), particularly during the SDC, although their contribution has been relatively small.³³ By contrast, the decline observed in technical

efficiency (the second-largest component) has exerted an increasingly negative impact on TFP growth. For the median bank, the negative contribution increased from –0.8% in 2006 to –1.9% in 2017. All in all, these results suggest that Total Factor Productivity in the euro area banking sector has decreased over the last decade, driven mainly by a decline in technical efficiency. This result is undesirable because euro area banks need to boost productivity in order to support much-needed profitability.

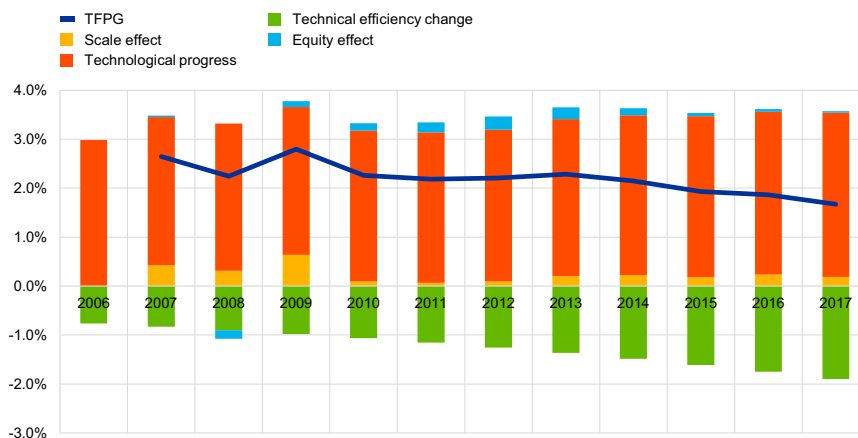
8 Conclusion

We employ an industrial organisation approach to study the performance of the banking system in the Eurozone over the period 2006–2017 as measured by Total Factor Productivity Growth (TFPG) and its components. The methodology is appealing because it allows the estimation of technical efficiency, technological progress and the equity and scale effects within the same econometric framework. To the knowledge of the authors, this is the first study that controls for unobserved bank heterogeneity and disentangles persistent and time-varying inefficiency in the euro area banking sector. This modelling choice is important because omitting bank heterogeneity and failing to distinguish persistent and time-varying inefficiency may lead to distorted and biased inefficiency estimates and because the set of policies needed in both cases is different.

We find that Total Factor Productivity in the euro area banking sector declined from above 2.6% in 2007 to below 1.7% in 2017, driven mainly by a decline in technical efficiency. Reverting this trend is difficult but desirable in a context where banks have been struggling to maintain profitability due to the low-interest rate environment and in a sector characterised by excess capacity and heavy cost structures. In this regard, boosting consolidation (cross-border and domestic) has the potential to foster economies of scale across the European banking sector, which according to our findings should result in cost savings, also for the largest banks. Such cost

³³ Duygun et al. (2015) find that the need to maintain acceptable equity capital ratios contributed to depress productivity growth in emerging market economies in the period preceding the global financial crisis.

Fig. 4 Total factor productivity growth and components (*median*). Notes: Total Factor Productivity growth (TFPG) is computed according to Eq. (1). Source: Author’s calculations based on BankFocus data



savings would allow banks to invest in large-scale digitalisation and information technology. Improving access to online banking could help to make greater reductions in the density of the branch network, helping to cut cost on a permanent basis. Technology can also help enhance operational efficiency by improving processes for loan origination and servicing and enable workloads to be shared better across employees in multiple locations. The digital transformation could also help banks to adapt their business models, opening the door to diversifying their products and therefore revenue sources, and to making it easier to compete in other jurisdictions, thereby increasing customers’ choices. These are all elements which are being considered and fostered by European policymakers.

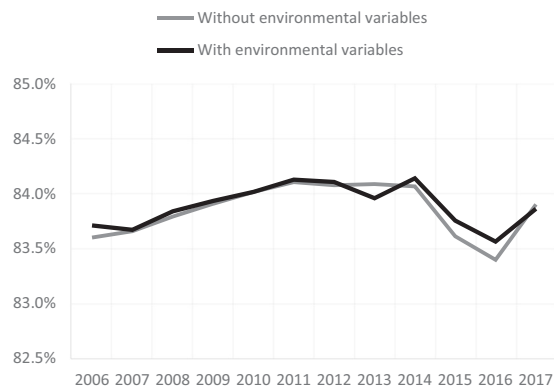
Acknowledgements The authors would like to thank, without implicating, two anonymous referees and an Associate Editor for very helpful comments and suggestions on an earlier version of this paper. We also thank participants in an internal seminar at the Macro Financial Linkages Division of the ECB, the Czech Economic Society and Slovak Economic Association Meeting (CES-SEAM) and a JVI Public Event. This paper should not be reported as representing the views of the European Central Bank (ECB), the Croatian National Bank (CNB) or the Joint Vienna Institute (JVI). The views expressed are those of the authors and do not necessarily reflect those of the ECB, the CNB and the JVI. The authors remain solely responsible for any remaining errors and omissions.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

9 Appendix: Efficiency scores when including environmental variables in Eq. (4)

This Appendix reports the overall efficiency estimates resulting from amending Eq. (4) to include environmental variables. The Chart below shows that differences between overall efficiency estimates with and without environmental variables are very small (to the third decimal) and that their

evolution is broadly similar.



References

Aigner D, Knox Lovell CA, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6:21–37

Altunbas Y, Goddard J, Molyneux P (1999) Technical change in banking. *Economics Letters* 64:215–221

Altunbas Y, Evans L, Molyneux P (2001) Bank ownership and efficiency. *Journal of Money, Credit and Banking* 33:926–954

Altunbas Y, Carbo S, Gardener EPM, Molyneux P (2007) Examining the Relationships between Capital, Risk and Efficiency in European Banking. *European Financial Management* 13:49–70

Babihuga R, Spaltro M (2014) Bank Funding Costs for International Banks, Working Papers, No.14/71. International Monetary Fund, Washington, D.C.

Badunenko O, Kumbhakar SC (2017) Economies of scale, technical change and persistent and time-varying cost efficiency in Indian banking: Do ownership, regulation and heterogeneity matter? *European Journal of Operational Research* 260:789–803

Balk B (2001) Scale Efficiency and Productivity Change. *Journal of Productivity Analysis* 15:159–183

Baltagi BH, Griffin JM (1988) A General Index of Technical Change. *Journal of Political Economy*, 1988 96:20–41

- Battese GE, Corra GS (1977) Estimation of a production frontier model: with application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics* 21:169–179
- Beccalli E, Anolli M, Borello G (2015) Are European banks too big? Evidence on economies of scale. *Journal of Banking & Finance* 58(C):232–246
- Berger AN (1993) Distribution-free estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions. *Journal of Productivity Analysis* 4:61–92
- Berger AN (1995) The profit-relationship in banking – tests of market-power and efficient-structure hypotheses. *Journal of Money Credit Banking* 27:405–431
- Berger AN, Hanweck TH, Humphrey DB (1987) Competitive viability in banking: Scale, scope, and product mix economies. *Journal of Monetary Economics* 20:501–520
- Berger AN, Mester LJ (1997) Inside the black box: what explains differences in the efficiencies of financial institutions. *Journal of Banking & Finance* 21:895–947
- Bonin JP, Hasan I, Wachtel P (2005) Bank performance, efficiency and ownership in transition countries. *Journal of Banking & Finance* 29:31–53
- Bos JWB, Schmiedel H (2007) Is there a single frontier in a single European banking market? *Journal of Banking & Finance* 31:2081–2102
- Boucinha M, Ribeiro N, Weyman-Jones T (2013) An assessment of Portuguese banks' efficiency and productivity towards euro area participation. *Journal of Productivity Analysis* 39:177–190
- Camanho AS, Dyson RG (2005) Cost efficiency, production and value-added models in the analysis of bank branch performance'. *Journal of the Operational Research Society* 56:483–494
- Casu B, Ferrari A, Girardone C, Wilson JOS (2016) Integration, productivity and technological spillovers: Evidence for eurozone banking industries. *European Journal of Operational Research* 255:971–983
- Colombi R, Kumbhakar SC, Martini G, Vittadini G (2014) Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis* 42:123–36
- Dijkstra, M. (2013): *Economies of scale and scope in the European banking sector 2002-2011, Working Paper, Amsterdam Center for Law & Economics, No. 2013-11. European Central Bank (ECB): "Cross-border bank consolidation in the euro area", Financial integration in Europe, ECB, May 2017.*
- Dissanayake R, Wu Y (2021), "Banking Efficiency Matters: Evidence from the Covid-19 Pandemic", *unpublished manuscript*.
- Duygun M, Shaban M, Sickles R, Weyman-Jones T (2015) How a regulatory capital requirement affects banks' productivity: an application to emerging economies. *Journal of Productivity Analysis* 44:237–248
- European Banking Authority (2018), Risk Assessment Questionnaire: Summary of Results, December.
- Farrell MJ (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society* 120:253–290
- Feng G, Serletis A (2009) Efficiency and productivity of the US banking industry, 1998–2005: Evidence from the Fourier cost function satisfying global regularity conditions. *Journal of Applied Econometrics* 24:105–138
- Filippini M, Greene W (2016) Persistent and transient productive inefficiency: a maximum simulated likelihood approach. *Journal of Productivity Analysis* 45:187–196
- Fiordelisi F, Molyneux P (2006) Shareholder Value in Banking. Palgrave Macmillan, UK
- Fiordelisi F, Marques-Ibanez D, Molyneux P (2011) Efficiency and risk in European banking. *Journal of Banking & Finance* 35:1315–1326
- Fungačova Z, Klein PO, and Weill L (2020), Persistent and transient inefficiency: Explaining the low efficiency of Chinese big banks, *China Economic Review*, 59.
- Goddard J, Liu H, Molyneux P, Wilson JOS (2013) Do Bank Profits Converge? *European Financial Management* 19:345–365
- Greene W (2005) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* 126:269–303
- Havranek T, Irsova Z, Lesanovska J (2016) Bank efficiency and interest rate pass-through: Evidence from Czech loan products. *Economic Modelling* 54:153–169
- Hughes J, and Mester L (1993), *Accounting for the demand for financial capital and risk-taking in bank cost functions, Working Paper, No. 93-17, Federal Reserve Bank of Philadelphia.*
- Hughes, J. and L. Mester (2008), *Efficiency in banking: theory, practice, and evidence, Working Papers, No. 08-1, Federal Reserve Bank of Philadelphia.*
- Hughes J, Mester L (2013) Who said large banks do not experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation* 22:559–585
- Hughes JP, Mester LJ, Moon C-G (2001) Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production. *Journal of Banking and Finance* 25:2169–2208
- Jonas MR, King SK (2008) Bank efficiency and the effectiveness of monetary policy. *Contemporary Economic Policy* 26:579–589
- Koetter M, Poghosyan T (2009) The identification of technology regimes in banking: Implications for the market power-fragility nexus. *Journal of Banking & Finance* 33:1413–1422
- Kumbhakar SC, Lien G, Hardaker JB (2014) Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis* 41:321–337
- Kumbhakar SC (1991) Estimation of technical inefficiency in panel data models with firm- and time-specific effects. *Economics Letters* 36:43–48
- Kumbhakar SC, Heshmati A (1995) Efficiency measurement in Swedish dairy farms: an application of rotating panel data, 1976–88. *American Journal of Agricultural Economics* 77:660–674
- Kumbhakar SC, Hjalmarsson L (1995) Labour-use efficiency in Swedish social insurance offices. *Journal of Applied Econometrics* 10:33–47
- Kumbhakar SC, Heshmanti A (1996) Technical changes and total productivity growth in Swedish manufacturing industries. *Econometric Review* 15:275–98
- Kumbhakar SC, Wang HJ, and Horncastle AP (2015), *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*, Cambridge University Press, May.
- Lensink R, Meesters A, Naaborg I (2008) Bank efficiency and foreign ownership: Do good institutions matter? *Journal of Banking & Finance* 32:834–844
- Lozano-Vivas A, Pasiouras F (2010) The impact of non-traditional activities on the estimation of bank efficiency: International evidence. *Journal of Banking & Finance* 34:1436–1449
- Markowitz H (1952) Portfolio Selection. *The Journal of Finance* 7:77–91
- Maudos J, Pastor JM, Perez F, Quesada J (2002) Cost and Profit Efficiency in European Banks. *Journal of International Financial Markets, Institutions and Money* 12:33–58
- Modigliani F, Miller MH (1958) The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review* 48:261–297
- Nițoi M, Spulbar C (2015) An Examination of Banks' Cost Efficiency in Central and Eastern Europe. *Procedia Economics and Finance* 22:544–551

- Psilaki M, Mamatzakis E (2017) What drives bank performance in transitions economies? The impact of reforms and regulations. *Research in International Business and Finance* 39:578–594
- Schlüter, T., R. Busch, T. Hartmann-Wendels and S. Sievers (2012), Determinants of the interest rate pass-through of banks: Evidence from German loan products, *Discussion Papers*, No. 26/2012, Deutsche Bundesbank.
- Shamshur A, Weill L (2019) Does bank efficiency influence the cost of credit? *Journal of Banking & Finance*, Elsevier 105:62–73
- Shen Z, Liao H, Weyman-Jones T (2009) Cost efficiency analysis in banking industries of ten Asian countries and regions. *Journal of Chinese Economic and Business Studies* 7:199–218
- Sealey C, Lindley J (1977) Inputs, outputs, and a theory of production and cost at depository financial institutions. *Journal of Finance* 32:1251–1266
- Tsionas EG, Kumbhakar SC (2014) Firm Heterogeneity, Persistent and Transient Technical Inefficiency: A Generalized True Random-Effects Model. *Journal of Applied Econometrics* 29:110–132