Economics and Business Letters
9(3), 270-278, 2020

Technical inefficiency and output scale in banking and industry

Gerasimos T. Soldatos*

American University of Athens, Athens, Greece

Received: 5 January 2020
Revised: 24 February 2020
Accepted: 26 February 2020

Abstract

What would the effect on bank leverage and industry production be if a weighted average of profit and inefficiency reduction were maximized in a perfectly competitive environment? The answer to this question by this note is that technical inefficiency unrelated to problem loans on the part of banks, would result in loan expansion, inducing, in turn, firms that rely heavily on bank borrowing, to expand their production accordingly as they can benefit from higher lending. Increased scale of production fosters economies of scale, which can compensate for the adverse effects of technical inefficiency on cost, reinforcing at the same time sales. Production can expand even further if firms maximize their profit by acknowledging their own technical inefficiency.

Keywords: technical inefficiency; weighted profit; increased leverage; increased real output
JEL Classification Codes: D21, D69, E52, G21

1. Introduction

Technical inefficiency as a scalar of a cost function coming out of a production function involving such an inefficiency (Khumbhakar 2006), reduces output if the agent being subject to this problem, disregards it. It follows that this, as any type of inefficiency, is undesirable on social welfare grounds. But, is the bearer of inefficiency the same unfavorable towards its presence? After all, other types of inefficiency like X-inefficiency and allocative inefficiency may be advantageous to banks and firms from the viewpoint of increased scale of production and market power. This paper argues that this is also the case with technical inefficiency when it is taken into account by the profit-maximizing calculus as a factor lowering "efficiency profit". To focus on technical inefficiency alone, a perfectly competitive framework is assumed. The argument made within this framework is simple: The only way a profit-maximizing agent can "defend against" the additional cost imposed by technical inefficiency, is to increase the

* E-mail: soldgera@yahoo.com.

Citation: Soldatos, G. T. (2020) Technical inefficiency and output scale in banking and industry, *Economics and Business Letters,* 9(3), 270-278.

DOI: 10.17811/ebl.9.3.2020.270-278

revenue side of the profit function by raising the volume of output provided for sale in the market exploiting at the same time economies of scale[1].

When ownership and management differ, the term "defense" connotes a weighted profit function in which one of the weights refers to the goal of cost-inefficiency reduction and the other weight to profit maximization; weights contained in manager's incentive contract. Under the customer(-consumer) sovereignty of perfect competition, allocative inefficiency is nil and cost inefficiency coincides with technical inefficiency, with the two terms being used interchangeably below. And, it is plausible to expect that the increased scale of output to which the weighted sum maximization leads, corroborates cost reductions and increased sales by means of economies of scale. In terms of the literature on production frontiers (Kontolaimou et al. 2012), the source of bank technical inefficiencies may include either input-based factors (like principal-agent problems), and/or bank-specific factors that affect bank output levels (like the internal organization of banks). Bank lending to producers is influenced accordingly; lending to firms that may too be subject to analogously classified technical inefficiencies. Also, following Klein (2013), non-performing loans are related to the state of macroeconomic activity rather than to cost efficiency considerations. In general, banking is being found by the banking efficiency literature to be subject to all sorts of inefficiency despite the heterogeneity in this literature (Aiello and Bonanno 2017). The emphasis analytically is on inefficiency that may be "dealt with" by upward shifts of the technical frontier; more so when as it appears empirically (Hughes et al. 2001), banks do increase the scale of output as a means of bringing about scale economies in order to counter adverse cost and revenue developments.

Indeed, as Rosenthal and Strange (2004) argue, scale economies shift the production function upwards, causing Hicks neutral effects. To get a glimpse of how the issue at hand is approached in relation to analytically similar literature, note that DeAngelo and Stulz (2015) attribute the secular increase in leverage to its social value, questioning thereby the validity of Modigliani-Miller (1958) theorem for banking, that is, the debt-equity neutrality. Alternative approaches to the matter that add leverage-related frictions − like moral hazard and taxes − to the neutrality baseline, suggest that increased leverage is induced by such inefficiencies, implying that regulators could reduce bank lending at no social cost (Admati and Hellwig 2013). In the next section, technical or the same, cost inefficiency if allocative inefficiency is nil, is not taken to be a leverage-related friction[2]. It is also supposed to be unrelated to problem loans. The argument then in the previous paragraph above makes clear that increased leverage should be expected under this formulation of inefficiency regardless of the issue of capital structure that pervades the standard literature on the topic under examination. Technical inefficiency results in the violation of Modigliani-Miller neutrality not because it is a friction to the neutrality, but because it would be profit-reducing if it were not reckoned by the bank.

This is shown formally in the next section where the duality between "profit maximization accounting for technical inefficiency" and "average profit maximization by management placing a weight on cost reduction" is also established. It is not plausible, of course, to expect an ownership-management split in the presence of perfect competition, but "the split" serves as a theoretical convenience. Section 3 argues that increased leverage to non-financial firms relying extensively on external finance, raises the output of these firms as documented empirically by Mirzaei and Moore (2019). Perfect competition does not rule out changes in output scale in the

[1] A combination of sales increase and price increase is expected to be the case under imperfectly competitive conditions. The subsequent difference in production and price relative to the values of these variables under perfect competition will be a measure of X-inefficiency, which is outside the realm of this paper.

[2] Modern cases of technical/cost inefficiency are according to CEO Buerger (2018) the complicated way online applications are administered, the costly single-product, single-channel pre-approval campaigns, and the seasonal and propensity-based targeting.

short- or medium-run. Even more output by firms will be the case if they themselves are subject to technical inefficiency and profit-maximizing takes this into account. Trying to neutralize this "extra" output will harm the ability of firms to exploit the benefits from increased lending by itself because it limits the scale of the output of the firm and hence, its access to credit. Intuitively, the kind of technical inefficiency expected to be pertaining predominantly to a perfectly competitive firm would be scale inefficiency. In any case, any policy attempt at reducing bank-inefficiency induced lending would entail sizeable social cost. Nevertheless, if such a policy is exercised, it is found that it will not be successful unless it is based on a set of relationships among reserve-requirements ratio, capital requirements percentage, and interbank rate. Section 4 concludes this note by expanding on policy matters.

2. The bank

Banks are assumed to be identical in a perfectly competitive environment. Let B is balance sheet constraint − the net position of the typical bank − L bank lending, D is deposits with the banks, and $r_L, r_D,$ and r, the lending, deposit, and interbank rates so that total identical-bank profit, Π_b, is:

$$\Pi_b = r_L L + rB - r_D D - aL^{\varepsilon} - \beta D^{\epsilon} \tag{1}$$

where coefficients $\varepsilon, \epsilon \in (0,1)$ while a and β are positive constants. The cost, C, is the sum of the two logistic-growth curves:$(aL^{\varepsilon} + \beta D^{\epsilon})$. If uC is the reduction in profit due to non-problem-loan technical inefficiency (Khumbhakar 2006), relation (1) becomes[3]:

$$\Pi_b^u = r_L L + rB - r_D D - aL^{\varepsilon} - \beta D^{\epsilon} - u(aL^{\varepsilon} + \beta D^{\epsilon}) \tag{2}$$

Letting $\rho \in (0,1)$ be the minimum reserve requirements and E equity, being equal to the product of L times the fraction $\theta \in (0,1)$ of the minimum capital requirements:

$$B = (1-\rho)D - L + \theta L \tag{3}$$

Inserting (3) in (2) yields:

$$\Pi_b^u = [r_L - r(1-\theta)]L - a(1+u)L^{\varepsilon} + [r(1-\rho) - r_D]D - \beta(1+u)D^{\epsilon} \tag{4}$$

with first-order maximization conditions:

$$\frac{\partial \Pi_b^u}{\partial L} = [r_L - r(1-\theta)] - \varepsilon a(1+u)L^{\varepsilon-1} = 0 \Rightarrow L = \left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \tag{5}$$

$$\frac{\partial \Pi_b^u}{\partial D} = r(1-\rho) - r_D - \epsilon\beta(1+u)D^{\epsilon-1} = 0 \Rightarrow D = \left[\frac{\epsilon\beta(1+u)}{r(1-\rho) - r_D}\right]^{\frac{1}{1-\epsilon}} \tag{6}$$

If $\varepsilon = \epsilon = 2$, the ratio L/D is not influenced by inefficiency[4].

Result 1: Given the assumptions surrounding (4), (5) states that inefficiency leads to increased lending relative to the case of efficiency when $u = 0$, in order to offset presumably through

[3] An alternative more complicated bank-profit function would be:

$$\ln \Pi_b = \ln(r_L L + rB - r_D D - aL^{\varepsilon} - \beta D^{\epsilon}) - \frac{aL^{\varepsilon} + \beta D^{\epsilon}}{r_L L + rB - r_D D - aL^{\varepsilon} - \beta D^{\epsilon}} u$$

Such a function presumes that $\Pi_b > 0$ and would only complicate the algebra unnecessarily because it is the qualitative character of inefficiency upon which this analysis focuses. The experience of the firm that makes it reconsider its view of profit may be ranging from having to cope with internal adjustment costs to learning by doing effects and the role of capital vintage (Tsekouras and Kounetas 2010).

[4] Note also that (5) and (6) will be positive *iff* $r_L - r(1-\theta) > 0$ and $r(1-\rho) - r_D > 0$ in the denominators of the fractions.

increased revenue the negative term $-u(aL^{\varepsilon}+\beta D^{\epsilon})$ in (2) or (4). This, in turn induces banks to demand more deposits as (6) indicates.

If banks were not aware of their (identical) technical inefficiency, L and D would be as in (5) and (6), respectively, but with $u=0$, and bank profit would not be maximized. Letting these suboptimal values of lending and deposits be λ and δ, correspondingly, $u\lambda$ more lending and $u\delta$ additional deposits would be needed to maximize profit. But, accounting for the presence of technical inefficiency, boosts the scale of production as this fosters economies of scale, which can compensate for the adverse effects of the inefficiency on revenue.

One way to rationalize u's presence in a profit function is through manager incentive-contract considerations. Such considerations arise in the presence of split between ownership, seeking only profit maximization, and management, seeking the attainment of non-profit targets too − targets like increased revenue or market share or cost reduction − as follows: Let the profit in (4) be identified with some contract or incentive scheme according to which a weighted average of profit and non-profit goal is maximized as in:

$$\gamma\Pi_b^u+(1-\gamma)F=\Pi_b^u-uC$$

where γ and $(1-\gamma)$ are the weights and F is the non-profit goal. Rearranging terms in this equality, yields that:

$$uC=(1-\gamma)(\Pi_b^u+F)$$

which when inserted back to (Π_b^u-uC), the following contract is obtained:

$$\Pi_b^u-uC=\Pi_b^u-(1-\gamma)(\Pi_b^u+F)=\gamma\Pi_b^u-(1-\gamma)F \qquad (4')$$

F is accompanied by the minus sign, which indeed should be the case if it was identified with the cost function, signifying a contract in which special emphasis on cost reductions is placed besides the profit maximization target (Cornes and Itaya 2016). Analytically, therefore, profit maximization accounting for technical inefficiency is dual to maximizing a weighted average of profit and cost reduction. Applying (4') to (4):

$$\begin{aligned}[r_L-r(1-\theta)]L-aL^{\varepsilon}+[r(1-\rho)-r_D]D-\beta D^{\epsilon}-(aL^{\varepsilon}+\beta D^{\epsilon})u\\ =\gamma\{[r_L-r(1-\theta)]L-aL^{\varepsilon}+[r(1-\rho)-r_D]D-\beta D^{\epsilon}\}\\ -(1-\gamma)(aL^{\varepsilon}+\beta D^{\epsilon})\end{aligned}$$

and zeroing the first derivatives of the right-hand side of this equality with respect to L and D, one obtains that:

$$\frac{\partial\Pi_b^u}{\partial L}=\gamma[r_L-r(1-\theta)]-\varepsilon aL^{\varepsilon-1}=0\Rightarrow L=\left\{\frac{\varepsilon a}{\gamma[r_L-r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \qquad (5')$$

$$\frac{\partial\Pi_b^u}{\partial D}=\gamma[r(1-\rho)-r_D]-\varepsilon\beta D^{\epsilon-1}=0\Rightarrow D=\left\{\frac{\varepsilon\beta}{\gamma[r(1-\rho)-r_D]}\right\}^{\frac{1}{1-\epsilon}} \qquad (6')$$

which when are equated to (5) and (6), respectively, yields that:

$$u=\frac{1-\gamma}{\gamma}$$

Result 2: Given the assumptions surrounding (4) and ownership-management difference, the technical inefficiency scalar, u, is equal to the ratio of the weight placed on the goal of cost reduction to the weight placed on profit maximization, *ceteris paribus*.

Result 1': Given technical inefficiency captured by uC, the weighted-profit maximizing weights given by Result 2, dictate increased lending relative to the case of efficiency when $u=0$ and $\gamma=1$.

Figure 1 in the Appendix illustrates the relationship between inefficiency scalar and weight on profit maximization; the higher the inefficiency the smaller this weight is.

Now, to find the response that r and θ should have to a change in u in order to leave loan provision unchanged, (5) is totally differentiated and equated to zero, resulting in the expression:

$$\frac{1}{(1+u)}du = \frac{(1-\theta)}{[r_L - r(1-\theta)]}dr + \frac{r}{[r_L - r(1-\theta)]}d\theta \tag{7}$$

But, since the deposits with the bank system should not be influenced by a change in u that leaves dL unchanged, setting $dD = 0$ in the total differential of (6), and equating to zero, yields:

$$\frac{1}{(1+u)}du = \frac{(1-\rho)}{[r(1-\rho) - r_D]}dr - \frac{r}{[r(1-\rho) - r_D]}d\rho \tag{8}$$

Result 3: The Modigliani-Miller theorem does not hold: One obtains by equating (7) and (8) that $dL = 0$ in response to du when *all* three policy variables, $r, \rho,$ and θ, respond to du by satisfying the following relationship:

$$\begin{aligned}[(1-\rho)r_L - 2(1-\rho)(1-\theta)r + (1-\theta)r_D]dr \\ = r[r(1-\rho) - r_D]d\theta + r[r_L - r(1-\theta)]d\rho\end{aligned} \tag{9}$$

given, of course, the assumptions surrounding (4) and Result 1.

The policymaker has surely a delicate task to perform. If, for instance, $\rho = 0$ as is the case in many economies nowadays, capital requirements and the interbank rate should be used in relation to each other as the following derivative from (9) exemplifies (after simplification of the fraction):

$$\frac{dr}{d\theta} = \frac{r}{[r(1-\rho) - r_D]} > 0$$

Nevertheless, since $du = -(1/\gamma^2)d\gamma$, the change in loan provision may be neutralized by simply asking banks to leave unchanged management contracts, because this would imply that $d\gamma = 0 \therefore du = 0$.

3. The firm

Next, let y be the total identical-firm product in a perfectly competitive market. y is produced by labor, N, and L, according to the Cobb-Douglas function: $y = N^{\upsilon}L^{\eta}$, where the sum of the constants υ and η captures returns to scale. In the absence of inefficiency in production, total firm profit, Π_f, when output price is p, and the wage rate is w, will be:

$$\Pi_f = pN^{\upsilon}L^{\eta} - wN - r_L L \tag{10}$$

with optimal input demands:

$$N' = \frac{p\upsilon y}{w} \tag{11}$$

$$L' = \frac{p\eta y}{r_L} \tag{12}$$

Equating (12) with (5) results in that at equilibrium:

$$\frac{p\eta y_u}{r_L} = \left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \Rightarrow y_u = \frac{r_L}{p\eta}\left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \tag{13}$$

where y_u is output when bank inefficiency is taken into account. That is, output production reflects bank inefficiency, which is as plausible to expect as that $\partial y_u/\partial u > 0$, given the ratio $(r_L/p\eta)$ that that $\partial L/\partial u > 0$. Setting $u = 0$ in (13), this output, y_u, is higher than it would be in the absence of bank inefficiency. Moreover, inserting (13) in (11) yields:

$$N' = \frac{vr_L}{w\eta}\left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \tag{11'}$$

Yet, inserting (13) in (12), one obtains that: $L' = L$.

Result 4: Given the assumptions surrounding (4) and (10), and Result 1, at equilibrium, the firm borrows the amount that optimizes inefficient-bank profit and output is increased by hiring more labor; indeed, $\partial N'/\partial u > 0$.

When firms are subject to technical inefficiency too, captured by v, (10) becomes:

$$\Pi_f^v = pN^v L^\eta - (1+v)wN - (1+v)r_L L \tag{14}$$

with optimal inputs demand:

$$N'' = \frac{pvy_v}{w(1+v)} \tag{15}$$

$$L'' = \frac{p\eta y_v}{r_L(1+v)} \tag{16}$$

At equilibrium, where (16) should be equal to (5):

$$\frac{p\eta y_{v,u}}{r_L(1+v)} = \left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \Rightarrow y_{v,u} = \frac{r_L(1+v)}{p\eta}\left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} = (1+v)y_u > y_u \tag{17}$$

Result 5: Given the assumptions surrounding (4), (10), and (14), and Results 1 and 3, (17) states that output under both inefficiencies, $y_{v,u}$, is higher than that under only bank inefficiency although setting (17) in (15), (11') comes up again while $L'' = L' = L$, because simply L is what maximizes bank profit.

It follows that producer (technical) inefficiency does not change borrowing per laborer at equilibrium. But, although employment and loan demand by the firm are influenced only by bank inefficiency, they result in more output in case there is also producer inefficiency. The reason is that increasing sales and thereby the revenue term in (14), can counteract the negative term $-v(wN + r_L L)$, which is not present in the absence of producer inefficiency. From the point of view of ownership-management difference, with μ being the weight on profit maximization and $(1-\mu)$ the weight on cost reduction in firm's management contract, $v = (1-\mu)/\mu$, and:

$$y_{v,u} = \frac{r_L}{\mu p\eta}\left\{\frac{\varepsilon a}{\gamma[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} \tag{17'}$$

In view of producer inefficiency, firm's management tries to raise revenue and reduce cost by attaining scale economies via increased production just as bank's management does in the presence of bank inefficiency. According to (17'), the greater the emphasis on cost reduction, i.e. the lower μ is, the higher real output will be given γ; (see also Figure 1 in the Appendix).

By how much a change in v and opposite thereby change in μ alter real output? Comparing (17) with (13), one concludes that when bank inefficiency is combined with output technical inefficiency:

$$\frac{\partial y_{v,u}}{\partial v} = \frac{r_L}{p\eta}\left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}} = y_u \tag{18}$$

Result 6: Given Result 4, a change, according to (18), in producer inefficiency alters output in the same direction by an amount which is equal to output in the presence of only bank inefficiency.

Result 6': A one percent increase in v will raise $y_{v,u}$ by $v/(1+v)$ percent or, the same, by $(1-\mu)$ percent, since:

$$\frac{\partial y_{v,u}}{\partial v}\frac{v}{y_{v,u}} = \frac{v}{1+v} = 1-\mu$$

That is, producer inefficiency is quite a powerful motive to increase production. This can also be seen by setting the total differential[5]:

$$dy_{v,u} \approx y_u dv + (1+v)dy_u \tag{19}$$

equal to zero, obtaining that:

$$\frac{dy_u}{dv} = -\frac{y_u}{1+v} \tag{20}$$

The derivative in (20) states that if $y_{v,u}$ is to remain constant in view of increased producer inefficiency, the inefficiency would reduce the output that would have been provided in the presence of bank inefficiency alone.

Result 7: Neutralizing the motive to produce more in response to producer inefficiency, makes this inefficiency hurt the ability of the firm to protect itself from the inefficiency introduced by the bank sector.

The best that might be done by firms if by "force", $dy_{v,u} = 0$ after $dv > 0$, would be to produce exactly y_u given that increased revenue from increased sales is the only way to counteract the negative influence of inefficiency on profit. This is at least what one concludes from (18). But, according to (20), producer inefficiency prevents firms from being able to do even this, because presumably this inefficiency consists mainly of scale inefficiency. $dv > 0$ signifies an increase in this inefficiency and hence, a decrease in the ability to borrow. These considerations are illustrated through Figure 2 in the Appendix as well.

Result 8: In general, a policy aiming at neutralizing bank-inefficiency induced increase in lending and/or its effects on real output, is undesirable, *ceteris paribus*.

4. Concluding remarks

Continuing to assume perfect competition, there is only one issue that could justify policy intervention in the context of this paper, namely the increased risk-taking involved necessarily when lending expands. Bank-inefficiency induced leverage may not be as extensive as to foster the perils of high-leverage finance capitalism but it does contribute to systemic fragility. If and when problems arise with this fragility, the conclusion of this analysis is that the problem should be tackled by both monetary and macroprudential policy as outlined, for example, by Shin (2015) or Bruno et al. (2017). Capital requirements, the interbank rate, and if there is a ratio of reserve requirements, are all interrelated. But, again, it is not the technical inefficiency per se that should be targeted; it is the problem of inappropriately high risk-taking.

Relaxing the assumption of perfect competition, it is possible that increased lending may prompt the emergence of imperfect competition. In reality, the banking system is imperfectly

[5] (20) derives as follows:

$$\frac{\partial y_{v,u}}{\partial u} = \frac{r_L}{p\eta}\left\{\frac{\varepsilon a(1+u)}{[r_L - r(1-\theta)]}\right\}^{\frac{1}{1-\varepsilon}}\frac{(1+v)}{(1-\varepsilon)(1+u)} = y_u\frac{(1+v)}{(1-\varepsilon)(1+u)} = \frac{\partial y_u}{\partial u}(1+v) \Rightarrow \frac{\partial y_{v,u}}{\partial y_u} \approx (1+v)$$

$$\therefore$$

$$dy_{v,u} = y_u dv + y_u\frac{(1+v)}{(1-\varepsilon)(1+u)}du$$

or, in view of derivative $\partial y_{v,u}/\partial u = y_u[(1+v)/(1-\varepsilon)(1+u)]$:

$$dy_{v,u} \approx y_u dv + (1+v)dy_u$$

competitive, and a different policy viewpoint derives once this is acknowledged, since one component of the X-inefficiency of this system, is the reduced leverage relative to the perfectly competitive one under efficiency. X-inefficiency counteracts the lending expansion induced by efforts to curb technical inefficiency, and some optimal mix between systemic fragility and the systemic stability accompanying increased loan market concentration is thought to exist (Smets 2014). Given allocative efficiency, X-inefficiency is related to market concentration and stability whereas technical or cost inefficiency is linked to systemic fragility. The optimal mix of stability and fragility is then an optimal mix of these two inefficiencies. It follows that any higher risk-taking relative to that at this optimum, should be attributed to reasons related to efficiency, ceteris paribus. From this point of view, a policy aiming at bank-manager incentive contracts in combination with a monetary-cum-macroprudential policy would be more effective than the latter policy alone.

Note that our results would not hold if there were no good coordination between loan officers and the personnel overseeing operation costs. But, this is not the case in reality, which poses the question whether cost inefficiency precedes or is preceded by loan expansion. Having assumed technical inefficiency unrelated to problem loans generated by loan overexpansion (and assuming away allocative efficiency issues), a causality running from cost inefficiency to loan growth has been presumed here despite the static character of the discussion. In other words, there is something other than problem loans that triggers cost inefficiency, and a manager comes to fix this situation by expanding lending, but not to the extent of engendering problem loans. That "something other" could be according to the literature on the subject, bad management, "skimping", or just bad luck (Berger and DeYoung 1997). In any case, what is important policy-wise according to this paper, is that a (new possibly) manager comes to take care of such problems and not monetary policy.

Acknowledgments

This paper has benefited from the useful comments and suggestions by two reviewers.

References

Admati, A., and Hellwig, M. (2013) *The bankers' new clothes*, Princeton University Press: Princeton, NJ.

Aiello, F., and Bonanno, G. (2017) On the sources of heterogeneity in banking efficiency literature, *Journal of Economic Surveys*, 32(1), 194-225.

Berger, A. N., and DeYoung, R. (1997) Problem loans and cost efficiency in commercial banks, *Journal of Banking and Finance*, 21(6), 849-870.

Bruno, V., Shim, I., and Shin, H.S. (2017) Comparative assessment of macroprudential policies, *Journal of Financial Stability*, 28, 183-202.

Buerger, D. (2018) Three inefficiencies financial institutions must correct immediately, *The Financial Brand*, December 19, 2018. https://thefinancialbrand.com/78480/digital-lending-solution-technologies-cx-experience/

Cornes, R., and Itaya, J. (2016) Alternative objectives in an oligopoly model: An aggregative game approach, CESifo Working Paper Series 6191. https://www.ifo.de/DocDL/cesifo1_wp6191.pdf

DeAngelo, H., and Stulz, R.M. (2015) Liquid-claim production, risk management, and bank capital structure: Why high leverage is optimal for banks, *Journal of Financial Economics*, 116(2), 219-236.

Hughes, J. P., Mester, L. J., and 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(12), 2169-2208.

Klein, N. (2013) Non-performing loans in CESEE: Determinants and macroeconomic performance, IMF Working Paper 72. https://ssrn.com/abstract=2247224.
Kontolaimou, A., Kounetas, K., Mourtos, I., and Tsekouras, K. (2012) Technology gaps in European banking: Put the blame on inputs or outputs?, *Economic Modelling*, 29, 1798-1808.
Khumbhakar, S. C. (2006) Specification and estimation of nonstandard profit functions, *Empirical Economics*, 31, 243-260.
Mirzaei, A., and Moore, T. (2019) Real effect of bank efficiency: Evidence from disaggregated manufacturing sectors, *Economica*, 86, 87-115.
Rosenthal, S. S., and Strange, W. C. (2004) *Evidence on the nature and sources of agglomeration economies*, in Henderson, J.V. and Thisse, J.F. (Eds.): Handbook of Regional and Urban Economics, Elsevier: Amsterdam, vol. 4, chap. 49, 2119-2171.
Shin, H. S. (2015) Macroprudential tools, their limits and their connection with monetary policy, BIS. https://pdfs.semanticscholar.org/45a2/e16c77da2d0572adf940b8a449de47ca14d7.pdf?_ga=2.9121078.2005111820.1577352218-1378868370.1576920930
Smets, F. (2014) Financial stability and monetary policy: How closely interlinked?, *International Journal of Central Banking*, 3(2), 263-300.
Tsekouras, K., and Kounetas, K. (2010) Are the energy efficient technologies efficient?, *Economic Modelling*, 27(1), 274-283

Appendix – Additional figures

Figure 1. Inefficiency scalar on the vertical axis, and weight on profit maximization along the horizontal axis. The higher the inefficiency the smaller this weight is.

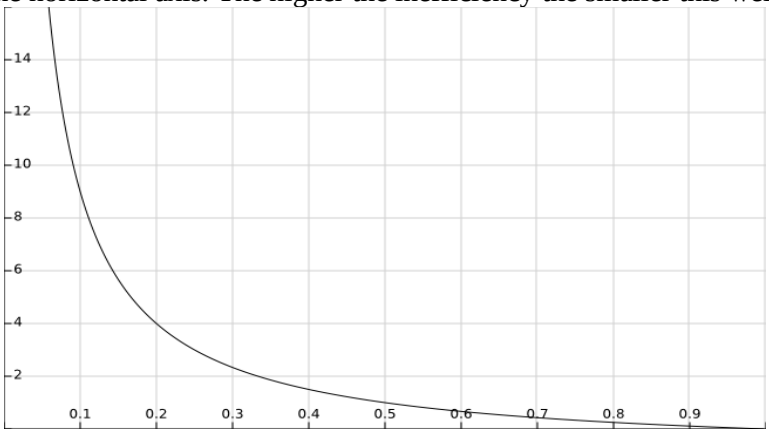

Figure 2. y1 and y2 are production curves. Given some inputs mix x, taking producers inefficiency into account leads to y2, reducing scale inefficiency from ΞΨ to ΘΛ.