Centre d'Etudes des Politiques Financières
Institut d'Etudes Politiques
47, avenue de la Forêt Noire
67000 STRASBOURG

Capacity-utilization and Efficiency in the European Banking Industry

Mohamed E. CHAFFAI
Université de Sfax

Michel DIETSCH[1]
Université Robert Schuman de Strasbourg

Preliminary version

October1999

[1] This research was supported by a grant of the ' Fondation de la Banque de France pour la Recherche'. Thanks to K. Kerstens and participants to the1998 Georgia Productivity workshop for helpful comments on an earlier version of the text.
Correspondence to Michel Dietsch, CEPF,
Institut d'Etudes Politiques de Strasbourg,
47, avenue de la Forêt Noire, 67000 Strasbourg, France.
Tel : (33)-3-88-41-77-08. E-mail : michel.dietsch@urs.u-strasbg.fr.

1. Introduction

The financial deregulation that occurred in the 1980s changed the nature of the competition in the European banking industry and provoked an acceleration of technological progress. Many observers have argued that these changes have created excess capacity. However, only a few studies have attempted to verify the existence of excess capacity in the banking sector (Frydl (1993), Dietsch (1994), Davis and Salo (1998)). The first objective of this paper is therefore to consider the measurement issue of capacity-utilization. We estimate whether there actually was excess capacity in the banking industry of three main European countries – France, Germany and Italy – during the 1990s.

To this aim, we use the capacity-utilization measure proposed by Berndt and Fuss (1986) and by Morisson (1985, 1988). Most of the papers that use these works have considered the case of the manufacturing industries. Here, we extend this methodology to address the specificity of the banking industry.

The second objective of the paper is to determine the relationship between capacity-utilization and cost efficiency in the banking sector. If banks are experiencing excess capacity, they will likely be cost inefficient, since short-run average costs are higher than long-run average costs. Following classical microeconomics standards, we expect a positive relationship between excess capacity and cost inefficiency in the long run. One reason is that the persistence of excess quasi-fixed inputs means that these inputs are underused in the long run, which induces excessive costs and technical inefficiency. Since the proportions of variable and fixed inputs may not be optimal with respect to market prices, it also produces allocative inefficiency. Hunter and Timme (1995) have explored the issue of short-run and long-run cost efficiency in the banking industry. However, they did not consider the capacity-utilization measurement issue. In this paper, we investigate the link between the two issues.

In the first section of this paper, we consider the conceptual aspects of excess capacity in the banking industry. The second section surveys methodologies for capacity-utilization measurement. Section three examines the model of the bank variable costs. Data and results are presented in the next section. The last section concludes.

2. The issue of excess capacity in the banking industry

2.1. Why does excess capacity exist ?

From an economic analysis of the concept, excess capacity is the capacity the firm « regrets » to have installed. Excess capacity refers to a situation where, in a firm or in an industry, the current capacity which results from the existing level of the fixed inputs is higher than the optimal capacity. The latter is the long run capacity the firm would install if it could start operating today, considering the « normal conditions » of the market.

Thus, capacities that were considered « normal » when they were installed are today considered « excessive ». In the banking industry, two main factors could explain the emergence of such excess capacity in recent times (Frydl, (1993), Dietsch, (1994), Davis and Salo, (1998)): technological and regulatory shocks, one the one hand, and structural decline of the demand for banking products, the so-called « decline of banking »(Gorton and Rosen, (1995)), on the other.

Competitive pressure should eliminate excess capacity in the short-run. However, excess capacity could persist if the costs induced by the elimination of excess capacity are higher than the gains coming from its elimination. In banking markets, « normal » conditions are not necessarily those of perfect competition. Technological or other natural barriers to entry resulting in imperfections in the industry could explain why it is costly to reduce excess capacity.

The most important of these barriers are sunk costs. Sunk costs are costs yielding long-term benefits to the firm that cannot be recovered when the firm leaves the industry. If sunk costs are low, the banking industry can reduce excess capacity. On the contrary, if sunk costs are high, excess capacity could persist over time. In other words, excess capacity could persist if a large part of the banking fixed costs are sunk costs.

In banking, sunk costs could arise from the existence of long-term customer relationships and from switching costs. First, banks and borrowers derive benefits from long-term customer relationships because the information extracted from these repeated relationships is private information, that gives the bank both market power on the customer and a competitive cost advantage over other banks in the market (Sharpe (1990)). However, due to the private nature of this information, the bank cannot sell it to other banks. This explains why investments in these relationships represent sunk costs for the bank. Second, there exist switching costs for customers in the banking markets. This characteristic could result in a strategic competition among banks in which they accept to sell their products with a discount in the short run in order to capture the customer over the long term. Thus, banks accept to reduce their short run profitability in order to extract market power from the banking relationships in the long run, what could induce sunk costs.

2.2. Two criteria of excess capacity

We can distinguish two criteria of excess capacity in the banking industry: decreasing costs and excessive risk. Concerning the criterion of decreasing cost, excess capacity refers to a situation where there are either too many firms in the industry or where the average level of output produced by existing firms is too low to experience the economies of scale that are typically generated by the fixed inputs. Consequently, the short-run costs are decreasing. So, decreasing costs are a sufficient condition for the existence of excess capacity. However, they are not a necessary condition.

Decreasing costs could also result from product differentiation. In the case of monopolistic competition, firms may be profitable even when producing a level of output below the level which minimizes costs. The reason is that the firm could raise the price if it produces less. In the banking industry, such a competition could occur due to the existence of market power in local banking markets. For instance, during the rate of deposit regulation period, banks were engaged in non-price competition that emphasized quality of service (under the form of better accessibility to banking services, better localization, and so on). Increased bank « brand » loyalty tends to result on the part of the banks' depositors, thereby allowing the bank to benefit from a decreased price elasticity of demand for its services. Thus, the decreasing costs result are in part due to the preferences of the clients: as depositors, they take advantage of better quality or accessibility of banking services. In such market conditions, this excess capacity is « voluntarily » maintained by the banks.

Consequently, « voluntary » excess capacity can exist in the banking industry. However, the scope of this paper includes only « involuntary » excess capacity, i.e., excess capacity resulting from a level of bank products below the equilibrium level in normal market conditions. In other words, bank capacity is excessive compared to the voluntary capacity desired by the banks clients. If there is excess capacity, bank profits are insufficient to completely cover fixed costs. The return on capital is below the long-term costs. The banks could still get positive profits (due to market power). However, profits would be higher if the banks could reduce excess capacity.

Concerning the criterion of excessive risk, the efficient capacity level for banks cannot be defined without taking into account risks. Thus, excess capacity could be defined as a situation of « excessive » risk, that is a situation in which the loans and investment policy of the bank produce an increase of risk without a corresponding increase of return. In other words, following a portfolio approach, there is excess capacity if the global risk level of a bank or of the banking industry increases relative to the minimum rate of return that insures long-run profitability. Observation shows that such an increase of risk relative to the « normal » return appeared during the second half of the 1980's in the U.S. (Gorton and Rosen (1995)) and in France (Dietsch (1994)).

In this financial approach, excess capacity refers to a situation where banks have an amount of funds higher than the amount they can lend without reducing profitability and increasing risk. However, in this paper, we will concentrate our attention to the criterion of decreasing costs.

3. The capacity-utilization measurement

The capacity-utilization rate can be defined as the ratio of actual output to capacity output. Therefore, the problem is to find a measure of capacity output. At the macro level, this measure is often derived from business surveys. Excess capacity is measured by the gap between current output and potential output, that is the level of output the firm declares it can produce without increasing its fixed inputs. At the micro level, following Berndt and Fuss (1986), capacity-utilization could be defined using the concept of a short-run cost function. In this approach, the potential or capacity output is the output at which the short-run average cost is tangent to the long-run average cost. Thus, the rate of capacity-utilization depends on the quantity of variable inputs applied to the quasi-fixed inputs through variation of demand.

Different econometric models have been developed in the literature in order to measure the capacity-utilization rate (Morisson (1985), Berndt and Hesse (1985)). These models have considered the case of a single product firm with one quasi-fixed input. Segerson and Squires (1990) have extended these models by considering multi-product firms with several quasi-fixed inputs.

Here, we briefly survey these earlier models. Denote $Y=(Y_1, Y_2,\ldots, Y_p)$, the vector of outputs produced by a bank, $X=(x_1, x_2,\ldots, x_p)$, the vector of its variable inputs, and $Z=(z_1, z_2 ,\ldots, z_k)$, the vector of the quasi-fixed inputs. Let $\omega = (\omega_{x1}, \omega_{x2},\ldots, \omega_{xp}; \omega_{z1},.., \omega_{zk}) = (\omega_x, \omega_z)$, the market input prices vector of these different inputs. Two different cost functions can be defined : the long-run total cost (LRTC) and the short-run total cost (SRTC), which give for each bank the minimum (total, variable) cost to produce the level of output Y. These functions are the following :

$$LRTC = C(Y, \omega_x, \omega_z) \quad (1)$$

$$SRTC = CV(Y, \omega_x , Z) + \omega_z' Z \quad (2)$$

Where CV(.) is the variable cost function. The difference between equations (1) and (2) is that the observed quasi-fixed inputs do not necessarily minimize SRTC. The producer have to choose the optimal level of quasi-fixed inputs which is the solution to the following program:

$$\nabla_Z \, CV(Y, \omega_x , Z) + \omega_z = 0 \quad (3)$$

If the firm is in a situation of excess or under capacity, the solutions of the short-run program (Z* in equation (3)) and the long-run program (the level of the actual quasi-fixed input Z) are different. In the case of one quasi-fixed input, for example, if $Z^* < Z$, then the firm has excess capacity.

From equation (3), we can also derive the shadow prices of the quasi-fixed inputs, according to the following formula :

$$\omega^*_z = - \nabla_Z \, CV(Y, \omega_x , Z) \quad (4)$$

If the firm minimizes costs, the shadow prices should be equal to the market prices of quasi-fixed inputs, given the observed output level Y and the observed variable input prices ω_x. However, if the shadow prices ω_z^* are lower than the observed prices ω_z, there is excess capacity.

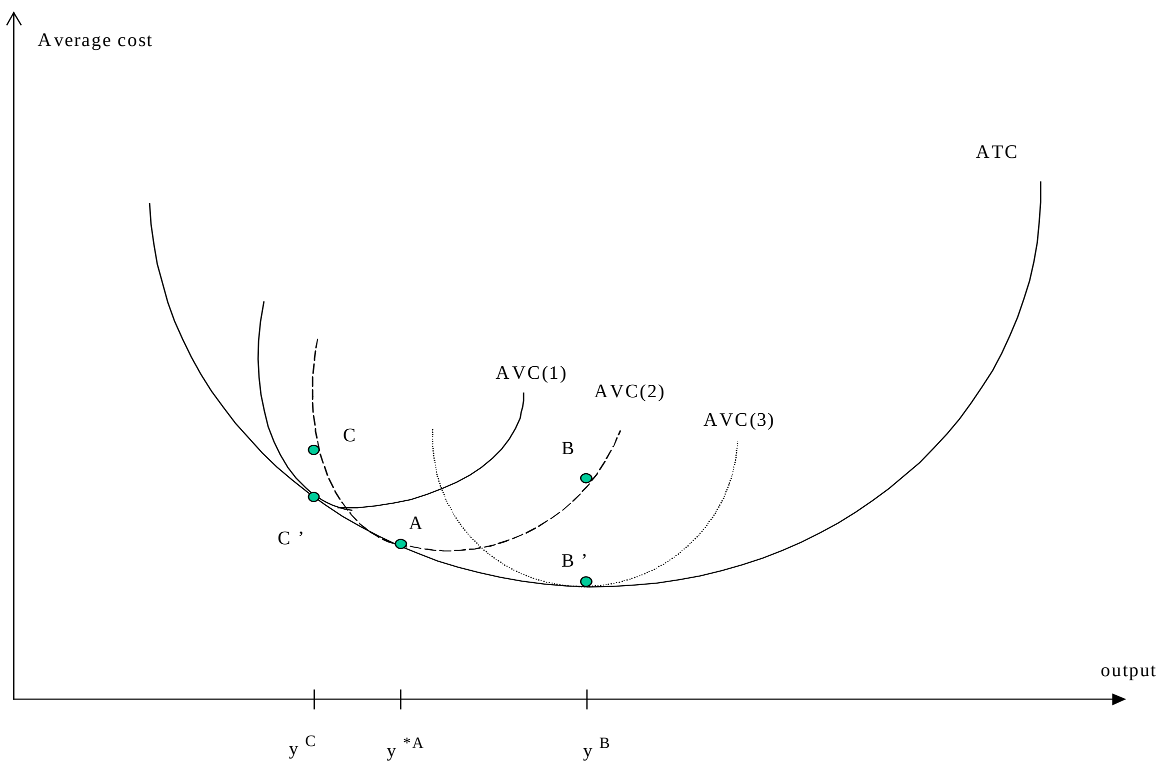
As an illustration, consider the case of a single output firm. A primal measure of capacity-utilization (CU_p) is given by the gap between the actual output Y and the potential or capacity output level Y*, i.e., the level that minimizes equation (3). This is the level at which the short-run average cost is tangent to the long-run average cost, i.e., $CU_p = Y/ Y^*$. This measure of capacity-utilization is illustrated in figure 1. In this figure, firm C does not minimize its short-run average cost. Its excess capacity is equal to $y^C/ y^{*A}<1$. The firm B is in the reverse situation of under capacity, which is measured by the ratio $y^B/ y^{*A}>1$. However, firm B could obtain full capacity if it would use more of the fixed input. By doing that, B could reduce average costs and would attain point B' in the figure. Firms A, C' and B' work at full capacity. Note that this definition of the capacity-utilization rate CU_P could only be used for the single output firm. In the case of multi-product technologies, Segerson and Squires (1990) propose three different primal measures of CU (total, ray, and partial CU measures). However, these measures require the introduction of very strong assumptions concerning the technology (homotheticity, separability, constant proportional decrease/increase of all outputs).

Capacity-utilization could also be measured using a dual approach, in which the capacity-utilization rate is the cost gap that exists when Y is not equal to Y*. This dual approach was proposed by Morisson (1985), who considered the single output technology case, and it has been extended later by Segerson and Squires (1990) for multi-product technologies. This dual measure is the cost gap between the potential cost of producing the actual output level Y assuming that average long-run cost is equal to average short-run cost (numerator of the equation (5) below), and the actual cost (denominator of (5)). The vector of shadow prices of quasi-fixed inputs is derived from equation (4). The dual measure capacity CU_d, is defined by the ratio of shadow total cost to actual total cost :

$$CU_d = \frac{TC^*}{TC} = \frac{CV(Y, \omega_x, Z) + \omega^{*\prime}_z Z}{CV(Y, \omega_x, Z) + \omega'_z Z} = 1 + \frac{(\omega^*_z - \omega_z)' Z}{\omega'_x X + \omega'_z Z} \quad (5)$$

If CU_d >(<1), we can conclude that the firm is in a situation of over (under) utilization of some fixed inputs, while if CU_d=1, shadow prices are equal to the quasi-fixed inputs market prices, and there is no excess or under capacity. Moreover, the interpretation of equation (5) is very easy in the case of a technology with one quasi-fixed input. In this case, if shadow price is greater (lower) than market price, there is no ambiguity in the result : CU_d is higher (lower) than 1. In the multi-quasi-fixed inputs case, by contrast, the relationship between shadow prices and market prices is more complex. For instance, if CU_d>(<1), it is only possible to say that at least one of the quasi-fixed inputs is over (under) used, which explains over(under)capacity.

The capacity-utilization rate could also be derived from the estimation of a short-run profit function (Segerson and Squires (1993)).



4. The model

In this section, we define a translog variable cost function with two quasi-fixed inputs and a composite error term for three European countries : France, Germany and Spain. This specification is the same as in Kaparakis, Miller and Noulas (1994) and Hunter and Timme (1995), where demand deposits and physical capital are the quasi-fixed inputs of the banks. The function can be written as follows:

$$
\begin{aligned}
Ln\,\mathrm{VC}_{\mathrm{it}} &= \alpha 0 + \sum_{\mathrm{j}} \alpha_j \,\mathrm{Ln}\,\omega_{\mathrm{jit}} + \sum_{\mathrm{j}} \sum_{j'} \alpha_{j\,\mathrm{j'}} \,\mathrm{Ln}\,\omega_{\mathrm{jit}} \,\mathrm{Ln}\,\omega_{\mathrm{j'it}} + \\
&\sum_{\mathrm{h}} \beta_h \,\mathrm{Ln}\,\mathrm{Y}_{\mathrm{hit}} + \sum_{\mathrm{h}} \sum_{h'} \beta_{hh'} \,\mathrm{Ln}\,\mathrm{Y}_{\mathrm{hit}} \,\mathrm{Ln}\,\mathrm{Y}_{\mathrm{h'it}} + \sum_{j} \sum_{h} \eta_{\mathrm{jh}} Ln\omega_{jit} LnY_{hit} \\
&\sum_{\mathrm{l}} \theta_{\mathrm{l}} Ln\,\mathrm{Z}_{\mathrm{lit}} + \sum_{\mathrm{l}} \sum_{\mathrm{l'}} \theta_{ll'} \,Ln\,\mathrm{Z}_{\mathrm{l\,it}} Ln\,\mathrm{Z}_{\mathrm{l'it}} + \sum_{\mathrm{j}} \sum_{\mathrm{l}} \delta_{jh} Ln\omega_{jit} LnZ_{lit} + \\
&\sum_{\mathrm{h}} \sum_{\mathrm{l}} \gamma_{\mathrm{hl}} \mathrm{Ln}\,\mathrm{Y}_{\mathrm{hit}} Ln\,\mathrm{Z}_{\mathrm{lit}} + \sum_{\mathrm{k}} \psi_{\mathrm{k}}\, D_{kit} + \mathrm{v}_{\mathrm{it}} + \mathrm{u}_{\mathrm{it}} \qquad (6)
\end{aligned}
$$

VC is the variable cost, which is the sum of labor costs, financial costs and other variable costs which are not related to the use of physical capital. The Y_i are the outputs produced by the bank : (1) saving and time deposits, (2) loans and (3) other earning assets. The ω_j are the unit prices of variable inputs : (1) price of labor, and (2) average interest paid on borrowed short-run and long-run funds. Z_1 and Z_2 are the quasi-fixed inputs. Z_1 represent demand deposits. In banking, the development of demand deposits (or core deposits) could be considered as an investment for the long run. Z_2 represent the fixed assets of the bank as a proxy measure of the stock of physical capital. v is the symmetric error term and u >0 is the asymmetric error term which represents cost inefficiency.

The variable cost function should verify some regularity conditions: homogeneity and symmetry, (these constraints are imposed on the model), as well as concavity in prices of variable inputs, and convexity in quasi-fixed inputs, (these conditions could not be imposed without loss of flexibility). More degrees of freedom could be obtained if we assume that the firms are minimizing costs and maximizing profit. This allows to introduce the following share equations in the model:

$$M_{jit} = \alpha_{j} + \sum_{j'} \alpha_{j'} \text{ Ln } \omega_{j'it} + \sum_{h} \eta_{jh} \text{ Ln}Y_{hit} + \sum_{l} \delta_{jl} \, Ln \, Z_{lit} \qquad (7)$$

$$S_{hit} = \beta_{h} + \sum_{h'} \beta_{hh'} \text{ Ln } Y_{h'it} + \sum_{j} \eta_{jh} \text{ Ln}\omega_{jit} + \sum_{l} \gamma_{lh} Ln \, Z_{lit} \qquad (8)$$

where the M_j are the variable cost shares of the variable inputs, while the S_h are the shares of the gross revenues coming from each output into the variable cost. In the cost function estimation, we retained one equation of revenue shares : the share of the loans' gross revenues in the variable costs.

To measure efficiency scores, we used the distribution free approach. This approach assumes that the differences owing to inefficiencies are stable over time while random errors are varying and tend to their average, zero, over time. As mentioned by Berger (1993), "good management maximises long-run profits by keeping costs relatively low over long periods of time, although costs may fluctuate from trend because of luck and measurement error". So by favouring a long-term perspective, managers of firms do not influence their efficiency from one year to the next one, the variations are all caused by chance or measurement errors.

To compute inefficiencies using DFA, each firm's inefficiency estimate is determined as the difference between the average residual of the firm and the average residual of the firm on the frontier. That is, the average of the annual residuals for each bank i is computed and serves as an estimate of $\ln x_i$ for that bank, given that the annual random error terms $\ln\omega_{it}$ tend to average out to zero over the period. These average residual of each bank i is used in the computation of efficiency. The efficiency score is given by the following equation: EFF = exp ($\ln x_{min} - \ln x_i$), where $\ln x_{min}$ is the minimum $\ln x_i$, i.e., the average residual for the bank with the lowest average cost residual, which is assumed to be the completely efficient bank. Therefore, EFF is an estimate of the predicted cost ratio for the most efficient bank to predicted costs for any bank. It is just like measuring inefficiency by the ratio of the predicted cost of the most efficient bank to predicted cost of each bank. Nevertheless, this efficiency measure is not completely correct if the random error terms $\ln\omega$ do not cancel each other out during the period. This error is likely to be larger for banks near the extremes of the average residual. These banks may have experienced good (bad) luck over the entire period. Consequently, the minimum average residual that serves as a benchmark for the calculation of the X-efficiency here could be overestimated. To treat this problem, we compute truncated measures of X-efficiency, where the value of the average residual of the qth ((1-q)th) percentile was attributed to each observation for which the value of the average residual is below (above) the qth ((1-q)th) quantile value. In this study, we use 5% truncation.

We have also estimated a long-run version of the cost model, where all inputs were considered as variable inputs. This model includes the total cost function - instead of the variable cost function - and the equations of the cost shares. The total cost inefficiency scores were derived according to the same methods described above.

Knowing these two cost inefficiency scores, we can investigate the relationship between capacity-utilization and cost efficiency. If the capacity-utilization rate is low in the short-run, while the long-run cost efficiency score is low, that means that the short-run cost inefficiency comes both from internal organization factors and from external demand factors. On the contrary, if CU is high, while the long-run cost efficiency is high, the external demand factors probably explain most of the short-run cost inefficiency.

5. The data and the results

This study covers the 1993-1997 period. We use unconsolidated accounting data for 1052 German, 130 French, and 178 Italian banks. Data come from the "Bankscope" database of BVD-IBCA. The banks included in the sample belong to three categories which differ by their legal statutes : corporate commercial banks, mutual banks and saving and loans banks. However, in Europe, these banks are submitted to the same regulation and they are in competition in the same markets. We excluded very small banks and also excluded specialized banks from the sample. All banks in the sample could be considered as “universal” banks, even if their size is quite small.

In what follows, we present first the evolution of the estimated capacity-utilization during the period 1993-1997. Then, we investigate the relationship between capacity-utilization and cost efficiency.

5.1. The evolution of capacity-utilization in three European banking industries over the 1993-1997 period

In order to get capacity-utilization rates year-by-year, we estimated the variable cost model for each year. The capacity-utilization rates were measured according to the dual method of Morisson (1985) and Segerson and Squires (1991) – see equation (5). These measures were obtained for each data point in the sample. Because market prices of the quasi-fixed inputs are not available, we used proxies for these market prices. We used the price of funds borrowed in the financial markets as a proxy for the market price of the physical capital, and we used an opportunity cost – the cost of interest bearing time deposits – as a proxy for the market price of the demand deposits. The use of this opportunity cost is justified as follows. The banks could substitute funds borrowed to their customers in the time deposit markets for demand deposits if they wanted to increase their activities. In that case, they will save the operating costs associated with the production of the transaction services ordinarily attached to the supply of demand deposits. Then, at equilibrium, the financial cost of these borrowed short-run funds should be equal to the non-financial cost of demand deposits. The sensitivity analysis of CU measurements to variations of these prices presented below justify this choice.

First of all, a likelihood ratio statistical test (not reported here) was performed to justify our choice of the two quasi-fixed inputs in the short-run cost function specification, that is demand deposits and physical capital. The null hypothesis that the variable cost function contains only one quasi-fixed input (physical capital or demand deposits) was tested against the alternative hypothesis that the variable cost function has the two quasi-fixed inputs. The likelihood ratio (LR) was compared to the table Chi-squared distribution with 7 degrees of freedom. This test rejected the model with one quasi-fixed input. We can conclude that in the short-run the banks of the three countries use demand deposits and physical capital as quasi-fixed inputs. Therefore, the results presented below were obtained by using the full short-run model (cost function and cost shares) with two quasi-fixed inputs.

Table 1 presents the evolution of the capacity-utilization over the period in the three countries. The capacity-utilization rates presented in this table are average annual values.

Table 1 : Average value of capacity-utilization in three European banking industries 1993-1997 (measures derived from a cost function)

	1993	*1994*	*1995*	*1996*	*1997*
France	0.836	0.861	0.866	0.873	0.831
Germany	0.850	0.854	0.853	0.871	0.857
Italy	0.822	0.782	0.806	0.823	0.733

The results show that there is a quite large excess capacity in the three European banking industries we consider during the period. However, this excess capacity seems to decrease slightly in France and Germany over the period. However, this trend tends to be confirmed. We observe that the excess capacity remains still high at the end of the period.

We also distinguished banks by size and we present capacity utilization estimates for the large banks and the small banks, in France and Germany. This distinction is necessary in order to take account for the heterogeneity of the domestic banking industries. Table 2 presents the results. It shows that, on average, over-capacity is higher in the small banks in Germany. However, the opposite result appears in France. While large German banks seem to have reduced over-capacity during the period, large French banks seem to have maintained it.

Table 2 : Average value of capacity-utilization by size in France and Germany 1993-1997 (measures derived from a cost function)

	1993	*1994*	*1995*	*1996*	*1997*
France Large banks	0.829	0.807	0.821	0.807	0.771
France small banks	0.841	0.912	0.918	0.958	0.894
Germany Large banks	0.829	0.855	0.877	0.906	0.938
Germany Small banks	0.879	0.861	0.832	0.863	0.830

One could propose some explanations of the high level of excess capacity at the beginning of the period. Before the banking deregulation process in the mid-1980s, the banking industry behaved like an oligopoly protected by public regulation. Because of a lack of price competition, banks competed at that time by opening new branches and expanding capacities. This has likely created over-branching and excess labor expenses. But, how to explain the persistence of excess capacity in the 1990s ? First, there was no consolidation of the banking industry in the three countries during the period. The acquisition of some very small banks by larger ones was not sufficient to contribute to the elimination of excess capacity. Second, strategic competition between banks do not induce banks to reduce operating costs. One of the main objectives of the banks was to increase their market shares. In the retail banking activities, that means that the capacity should not be reduced, even if the increase of price competition provoked a substantial decrease of the interest margins. Third, the economic recession of the mid-1990s could explain the stability of the capacity-utilization rate during the period under study.

As mentioned before, capacity-utilization rates could be sensitive to market prices of the quasi-fixed inputs (shadow prices of the physical capital and of demand deposits), as shown by equation (5). To investigate this possibility, we have measured CU rates while varying the market prices of the quasi-fixed inputs. We considered variations of the two prices between 3 % and 8 %. The CU rates were calculated for each bank. The results (not reported here) show that the CU rate is not too sensitive to the variations in the market prices. Indeed, most of the average capacity-utilization rates obtained by using this methodology were quite close to the capacity-utilization rate of our sample for each country. So, in the realistic ranges of market prices we consider, there still remains excess capacity in the three European banking industries we study.

5.2. The relationship between cost efficiency and capacity utilization

In the short run, over-capacity does not necessarily imply cost inefficiency. Excess capacity means that the quasi-fixed inputs are in excess and could be used more intensively, while short-run cost inefficiency means that the variable inputs are misused or misallocated, which results in an increase of costs. In other terms, in the short-run, a bank could be cost efficient even if it maintains excess capacity. However, in the long-run, over-capacity implies cost inefficiency. The excess of fixed inputs implies technical and allocative cost inefficiencies.

Table 3 : The relationship between short-run cost efficiency and capacity-utilization

	France	*Germany*	*Italy*
Average Capacity Utilization Rate during the 1993-1997 period	0.841	0.859	0.806
Average Efficiency Score during the 1993-1997 period *	0.913	0.900	0.804
Correlation between Efficiency scores and CU rates	-0.04	-0.05	-0.114

*The efficiency score is measured by using the variable (short-run) cost efficiency frontier

Table 3 presents the measures of average capacity-utilization rates CU over the 1993-1997 period, and the variable (short-run) cost efficiency scores EFF. The table presents also the correlation coefficients between variable cost efficiency and the CU rates in each countries.

These results permit us to reject the hypothesis that the banks are in a situation of full capacity. We can also reject the assumption that the banks are operating on their efficient short-run cost frontiers: on average, the French banks were inefficient by 9 % during the period, the German banks by 10 %, and the Italian banks by 20 %. Finally, there is no significant relationship between short-run cost efficiency and capacity utilization. The correlation coefficient between the two rates is very low in each country.

Table 4 presents the measures of the long-run cost efficiency scores EFF. In the long-run cost function, all inputs are variable inputs. The table presents also the correlation coefficients between long-run cost efficiency and the CU rates in each country.

Table 4 : The relationship between long-run cost efficiency and capacity-utilization

	France	*Germany*	*Italy*
Average Capacity Utilization Rate during the 1993-1997 period	0.841	0.859	0.806
Average Efficiency Score during the 1993-1997 period *	0.917	0.881	0.882
Correlation between long-run efficiency scores and CU rates	0.301	0.100	0.043

*The efficiency score is measured by using the long-run cost efficiency frontier

Notice that the short-run and long-run cost efficiency scores are also very close one another. Moreover, there is a very strong correlation between short and long run cost efficiency (0.84)[2]. That means that the banks that are cost efficient in the short-run have on average good chances to be also cost efficient in the long-run.

There is a significant correlation between capacity-utilization rates and cost efficiency. So, excess capacity seems to result in an increase of long-run costs. That confirms that over-capacity could be costly in the long run. However, the correlation coefficient is quite low in two countries.

If excess capacity persists in the long-run, one reason could be that these capacities are profitable. To test this hypothesis, we have estimated profit efficiency by using a profit function instead of a cost function.

5.3. The measurement of capacity utilization by using a profit function

The CU measurement could also be obtained in a alternative way by considering dual restricted profit functions. In this case, capacity utilization measurement take into account the demand fluctuations and the different choices of the output levels. We estimated a restricted standard[3] translog profit function, with one output prices (average return on loans and securities portfolio), two input prices, (average price of labor and of financial capital), and

[2] Hunter and Timme (1995) report a similar result : they find a coefficient correlation of 0.50 / 0.55 for U.S. banks.

[3] The non standard restricted profit function was not regular, in particular the shadow prices for the quasi fixed inputs were negative for most of the observations in the data set. The standard restricted profit function was regular.

two quasi-fixed inputs (demand deposits and physical capital). The endogenous variable in the profit function is total revenue minus variable costs. We also add the usual share equations to the model, which is estimated in the same manner as the variable cost model.

In this case, the measure of capacity utilization CU_P is obtained by comparing the actual profit of each bank to its shadow profit. The latter is obtained by replacing the fixed costs by their shadow cost values, where the shadow prices are obtained as the solution of the derivative of the variable profit function with respect to each quasi fixed input. According to Segerson and Squires (1993), the capacity utilization rate CU_P is equal to :

$$CU_P = \frac{PROF}{PROF^*} = \frac{Revenue - \omega_x^{'} X - \omega_z^{'} Z}{Revenue - \omega_x^{'} X - \omega_z^{*'} Z} \quad (9)$$

Table 5 presents the results. They show that, measured by using a non standard profit function, the excess capacity is still large in Italy (around 20%). However, in France and in Germany, results show a slight under-capacity. That could explain why banks maintain costly capacities in these two countries in the short-run.

Table 5 : Average value of capacity-utilization measured by using a profit function in three European banking industries 1993-1997

	1993	*1994*	*1995*	*1996*	*1997*
France	1.012	0.985	0.986	1.067	1.057
Germany	1.011	1.054	1.041	1.056	1.049
Italy	0.834	0.833	0.841	0.813	0.791

6. Conclusion

In this paper, we provide measures of capacity-utilization in three European banking industries during the period 1993-1997. These measures were derived from the estimation of the short-run variable cost frontier and from a short-run profit function with two quasi-fixed inputs, the fixed assets and the demand deposits. The results show that a significant excess capacity of approximately 15% to 20%, on average, continues to exist at the end of the period in the three European countries, France, Germany and Italy.

We also found no clear correlation between capacity-utilization and cost efficiency in the short-run. Over-capacity does not clearly result in cost inefficiency in the short-run. However, the relationship between cost efficiency and capacity utilization is clearer in the long-run. On average, excess capacity explains a significant part of the long-run cost inefficiency of the French and German banking industries. One explanation of the persistence of over-capacity and of the absence of a clear relationship between cost efficiency and the rate of capacity utilization in the short-run could come from the profit function. We observed that, in France and in Germany, there is no excess capacity in terms of profit.

Finally, cost or profit efficiency is a mix of technical efficiency and allocative efficiency. It could be interesting to investigate the relationship between these two forms of efficiency and excess capacity in the banking industry in the future. Moreover, it should be interesting to extend the model to distinguish more precisely the effect of fluctuations in the demand for banking products on the capacity-utilization rate.

References

BERGER A. (1993): "Distribution-Free Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distributional Assumptions", *Journal of Productivity Analysis* 4, 261-292.

Berndt E. R., and Hesse D. M. (1986) "Measuring and assessing capacity-utilization in the manufacturing sectors of nine OECD countries", *European Economic Review*, vol. 30, 961-989.

Davis P. and Salo (1998) " working paper, LSE-FMG, october 1998

Dietsch M (1994) "Les surcapacités bancaires en France", *Revue d'Economie Financière*, mars.

Eakin K. B., McMillen D. P., and Buono M. J. (1990) "Constructing confidence intervals using the bootstrap: An application to the multi-product cost function", *Review of Economics and Statistics*, 339-344.

Hulten C. R. (1986) "Productivity change, capacity utilization, and the sources of efficiency growth". *Journal of Econometrics*, vol. 33, 31-50.

Hunter W. C., and Timme S. G. (1995) "Core deposits and physical capital: A reexamination of bank scale economies and efficiency with quasi-fixed inputs". *Journal of Money Credit and Banking*, vol. 27, n°1, 165-184.

Jondrow J., Lovell C.A.K., Materov I. S., and Schmidt P.(1982) "On the estimation of technical inefficiency in the stochastic frontier production function model" *Journal of Econometrics*, 19, 233-238

Karapakis E. I., Miller S. M., and Noulas A. G.,(1994) "Short-run cost inefficiency of commercial banks: A flexible stochastic frontier approach", *Journal of Money, Credit and Banking*, vol. 26, n°4, 875-892.

Morisson C. (1985) "Primal and dual capacity utilization: An application to productivity measurement in the U.S. automobile industry" *Journal of Business and Economic Statistics* 53, 312-332

Morisson C. (1988) "Quasi-fixed inputs in United States and Japanese manufacturing: A generalized Leontief cost function approach" *Review of Economics and Statistics*, 70, 275-287

Morisson C. (1992) "Unraveling the productivity growth slowdown in the United States, Canada and Japan: The effects of subequilibrium, scale economies and markups" *Review of Economics and Statistics*, 3, 381-393

Schmidt P., and Sickles R. (1984) "Production frontiers and panel data" *Journal of Business and Economic Statistics*, 2, 367-374

Segerson K, and Squires D. (1990) "On the measurement of economic capacity-utilization for multi-product industries", *Journal of Econometrics*, vol. 44, 347-361

Segerson K, and Squires D. (1993) "Capacity-utilization under regulatory constraints". *Review of Economics and Statistics*, 76-85