

COST EFFICIENCY, SCALE ELASTICITY AND SCALE ECONOMIES IN ARAB BANKING

Idries Al-Jarrah, Philip Molyneux

Abstract

This study investigates Cost Efficiency, Scale Elasticity and Scale Economies of the Jordanian, Egyptian, Saudi Arabian and Bahraini banking systems. Our sample comprises information on 82 banks operating in the countries under study over 1992-2000. The stochastic frontier and Fourier-flexible form are used to estimate cost efficiency, scale elasticity and scale efficiency levels. The cost efficiency averaged around 95% over the 1992-2000 period. Islamic banks are found to be the most cost efficient while investment banks are the least. The cost scale elasticity estimates reveal diseconomies of around five percent and the cost scale inefficiency estimates suggest that banks are 65% scale efficient.

Key words: Efficiency, Cost Efficiency, Scale Economies, Scale Efficiency, Arab Banking, The Stochastic Frontier, the translog, the Fourier-flexible, Jordan, Saudi Arabia, Bahrain, Egypt, Bank Performance.

JEL Classification: G.

1. Summary

This study investigates Cost Efficiency, Scale Elasticity and Scale Economies of the Jordanian, Egyptian, Saudi Arabian and Bahraini banking systems. The empirical evidence on bank efficiency in these markets aims to highlight the features associated with the role of economic and financial reforms that have taken place in these countries over the past decade.

Our sample comprises information on 82 banks operating in Jordan, Egypt, Saudi Arabia and Bahrain over the 1992-2000 period. We use the stochastic frontier and Fourier-flexible form to estimate cost efficiency, scale elasticity and scale efficiency levels in these banking systems. The sample size represents 78% of the banking sector of Jordan, just under 90% of the Egyptian banking sector, 63% of that of Saudi Arabia and over 50% of the banking sector of Bahrain.

To derive efficiency levels, we employ cost efficiency concept using a number of different measurement methods (including the stochastic frontier approach, specification of the Fourier-flexible functional form versus the translog form, and inclusion of a banks' asset quality and financial capital in a number of different ways) to a single data set. In choosing the 'preferred' cost model to estimate efficiency levels, we follow various contemporary methodologies that use a variety of hypotheses tests to arrive at preferred model specifications. Our preferred model is the Fourier-truncated form that excludes the control variables (capital adequacy, asset quality and the time trend) but includes all the environmental variables.

The cost efficiency averaged around 95%, based on our preferred model, over the 1992-2000 period. Islamic banks are found to be the most cost efficient while investment banks are the least. This result perhaps reveals the fact that the cost of funds for Islamic banks is relatively cheaper than the cost of funds for other financial institutions. Large banks, in assets terms, appear to be relatively more cost profit efficient. This possibly signals the ability of large banks to utilise more efficient technology with less cost, the ability of these banks to introduce more specialised staff for the most profitable activities and the ability of these banks to provide (presumably) better quality outputs for which they can charge higher prices. Geographically, Bahrain is the most cost efficient while Jordan is the least.

Based on the estimated preferred model, we also report scale elasticity and scale efficiency measures for the banks under study. The cost scale elasticity estimates reveal diseconomies of around five percent and the cost scale inefficiency estimates also suggest that banks are 65% scale efficient. Islamic and commercial banks are again found to be the most cost scale efficient. Large banks are also generally found to be more efficient than smaller institutions. In addition, geographically, Saudi Arabian and Egyptian banks seem to be the most cost scale efficient.

The derived efficiency levels for the banks operating in the countries under study, however, provide little evidence to suggest that the economic and financial reforms undertaken in Jordan, Egypt, Saudi Arabia and Bahrain, over the last decade, have had a noticeable impact on improvement in banking sector efficiency. The main policy recommendation from this study, therefore, is that these countries need to continue the reform process in order to enhance financial sector performance.

2. Methodology: Measures of Efficiency and Productivity

The stochastic frontier, with the Fourier-flexible functional form, is the main methodology to be employed to derive efficiency measures in the countries under study. While the translog functional form has been probably the most widely utilised to derive efficiency estimates, the Fourier-flexible has received more focus in the recent efficiency literature.

The Fourier-flexible functional form

The stochastic cost model for a sample of N firms can be written as:

$$\ln TC_i = \ln TC(y_i, w_i, z_i; B) + u_i + v_i, i=1, \dots, N,$$

where TC_i is observed cost of bank i , y_i is the vector of output levels and w_i is the vector of input prices for bank i . z_i represents a vector of control variables which in the case of our estimates includes the quality of bank's output (q_i), the level of its financial capital (k_i) and the time trend (T_i). B is a vector of parameters, v_i is a two-sided error term representing the statistical noise (assumed to be independently and identically distributed and have a normal distribution with mean 0 and variance σ_v^2).

u_i are non-negative random variables that account for technical inefficiency. In case of Battese and Coelli (1995) model, u_i are assumed to be independently distributed as truncations at zero of the $N(m_i, \sigma_u^2)$ distribution; where $m_i = \delta_i d$, where δ_i is a set of environmental variables (defined in the previous section) which are employed to control for firm's specific factors that may contribute to explain the differences in the efficiency estimates, and d is a vector of parameters to be estimated. In case of Battese and Coelli (1992) model, u_i are assumed to be iid as truncations at zero of the $N(\mu_i, \sigma_u^2)$ distribution. The translog functional form for the cost frontier is specified as:

$$\begin{aligned} \ln(C/w_3) = & \alpha + \sum_{i=1}^2 B_i \ln(w_i/w_3) + \sum_{k=1}^3 \gamma_k \ln y_k + \sum_{r=1}^3 \psi_r \ln z_r \\ & + \frac{1}{2} \left[\sum_{i=1}^2 \sum_{j=1}^2 B_{ij} \ln(w_i/w_3) \ln(w_j/w_3) \right] + \frac{1}{2} \left[\sum_{k=1}^3 \sum_{m=1}^3 \gamma_{km} \ln y_k \ln y_m \right] \\ & + \frac{1}{2} \left[\sum_{r=1}^3 \sum_{s=1}^3 \psi_{rs} \ln z_r \ln z_s \right] + \sum_{i=1}^2 \sum_{k=1}^3 \eta_{ik} \ln(w_i/w_3) \ln(y_k) \\ & + \sum_{i=1}^2 \sum_{r=1}^3 \rho_{ir} \ln(w_i/w_3) \ln(z_r) + \sum_{k=1}^3 \sum_{r=1}^3 \tau_{kr} \ln y_k \ln z_r + u_{it} + v_{it} \end{aligned}$$

By augmenting the previous translog form by Fourier trigonometric terms, we get the Fourier-flexible functional form written as:

$$\begin{aligned}
\ln(C/w_3) &= \alpha + \sum_{i=1}^2 B_i \ln(w_i/w_3) + \sum_{k=1}^3 \gamma_k \ln y_k + \sum_{r=1}^3 \psi_r \ln z_r \\
&+ \frac{1}{2} \left[\sum_{j=1}^2 \sum_{i=1}^2 B_{ij} \ln(w_i/w_3) \ln(w_j/w_3) \right] + \left[\sum_{k=1}^3 \sum_{m=1}^3 \gamma_{km} \ln y_k \ln y_m \right] \\
&+ \frac{1}{2} \left[\sum_{r=1}^3 \sum_{s=1}^3 \psi_{rs} \ln z_r \ln z_s \right] + \sum_{i=1}^2 \sum_{k=1}^3 \eta_{ik} \ln(w_i/w_3) \ln(y_k) \\
&+ \sum_{i=1}^2 \sum_{r=1}^3 \rho_{ir} \ln(w_i/w_3) \ln(z_r) + \sum_{k=1}^3 \sum_{r=1}^3 \tau_{kr} \ln y_k \ln z_r \\
&+ \sum_{n=1}^8 [\phi_n \cos(x_n) + w_n \sin(x_n)] + \\
&\sum_{n=1}^8 \sum_{q=n}^8 [\phi_{nq} \cos(x_n + x_q) + w_{nq} \sin(x_n + x_q)] \\
&+ \sum_{n=1}^8 [\phi_{nmn} \cos(x_n + x_n + x_n) + w_{nmn} \sin(x_n + x_n + x_n)] + u_{it} + v_{it},
\end{aligned}$$

where $\ln C$ is the natural logarithm of total costs (operating and financial); $\ln y_i$ is the natural logarithm of bank outputs (i.e. loans, securities, off-balance sheet items); $\ln w_i$ is the natural logarithm of i th input prices (i.e. wage rate, interest rate and physical capital price); the x_n terms, $n=1, \dots, 8$ are rescaled values of the $\ln(w_i/w_3)$, $i=1, 2$, $\ln(y_k)$, $k=1, 2, 3$, and $\ln(z_r)$, $r=1, 2, 3$, such that each of the x_n span the interval $[0, 2\pi]$, and π refers to the number of radians here (not profits), and $\alpha, \beta, \gamma, \psi, \rho, \tau, \eta, d, \omega, \phi$ and t are coefficients to be estimated.

Since the duality theorem requires that the cost function be linearly homogeneous in input prices and continuity requires that the second order parameters are symmetric, the following restrictions apply to the parameters of the cost function in the equation above: $\sum_{i=1}^3 \beta_i = 1$;

$\sum_{i=1}^3 B_{ij} = 0$; $\sum_{i=1}^3 \eta_{ij} = 0$; $\sum_{i=1}^n \rho_{ij} = 0$; for all j . Moreover, the second order parameters of the cost function must be symmetric, that is, $B_{ij} = B_{ji}$ and $\eta_{ik} = \eta_{ki}$, for all i, k . The scaled log-output quantities; x_i are calculated as in Berger and Mester (1997) by cutting 10% off each end of the $[0, 2\pi]$ interval so that the z_i span $[0.1 \times 2\pi, .9 \times 2\pi]$ to reduce approximation problems near endpoints. The formula for z_i is $[0, 2\pi - \mu \times a + \mu \times \text{variable}]$, where $[a, b]$ is the range of the variable being transformed, and $\mu \equiv (0.9 \times 2\pi - 0.1 \times 2\pi / (9b - a))$. This study applies Fourier terms only for the outputs, leaving the input price effects to be defined entirely by the translog terms, following Berger and Mester (1997). The primary aim is to maintain the limited number of Fourier terms for describing the scale and inefficiency measures associated with differences in bank size. Moreover, the usual input price homogeneity restrictions can be imposed on logarithmic price terms, whereas they cannot be easily imposed on the trigonometric terms.

The maximum-likelihood estimates for the parameters in the Fourier-flexible stochastic frontier for Cost, Standard and Alternative profit efficiency functions; that includes efficiency correlates, are estimated using the computer program FRONTIER Version 4.0 (see Coelli, 1996). This computer program uses three steps to obtain the maximum likelihood estimates. The first step involves obtaining ordinary least squares (OLS) estimates of the equation. These estimates are unbiased because of the non-zero expectation of u_{it} . The second step involves evaluating the log-likelihood function for a number of values of γ between zero and one. During this procedure, d_i are set to zero

and the values of B_0 and σ^2 are adjusted according to the corrected ordinary least squares formulae for the half-normal model. The estimates corresponding to the largest log-likelihood value in this second step are used as starting values in the iterative maximisation procedure in the third and final part of the estimation procedure.

Calculation of within-sample scale elasticities

This study also estimates scale elasticities for the banks under study. Scale elasticity for the cost function (i.e., scale economies) refers to the proportional increase in cost resulting from a small proportional increase in the level of output (the elasticity of total cost with respect to output). Within the sample scale elasticities are calculated as in Mester (1996) and Altunbas et al. (1998) and are evaluated at the mean output, input price, asset quality and financial capital levels for the respective size quartiles. The degree of scale elasticities is given by the sum of individual cost elasticities. For the case of FF cost function, the measure of overall economies of scale (SE) is given by the following cost elasticity by differentiating the cost function in the above equation with respect to output;

This gives us:

$$SE = \sum_{i=1}^3 \frac{\partial \ln TC}{\partial \ln y_i} = \sum_{k=1}^3 \gamma_k + \sum_{k=1}^3 \sum_{m=1}^3 \gamma_{km} \ln y_m + \sum_{i=1}^2 \sum_{k=1}^3 \eta_{ik} \ln(w_i / w_3) + \sum_{k=1}^3 \sum_{r=1}^3 \tau_{kr} \ln z_r + \sum_{n=1}^8 [-\phi_n \sin(x_n) + \omega_n \cos(x_n)] + \sum_{n=1}^8 \sum_{q=n}^8 [-\phi_{nq} \sin(x_n + x_q) + \omega_{nq} \cos(x_n + x_q)] + \sum_{n=1}^8 [-\phi_{nm} \sin(x_n + x_n + x_n) + \omega_{nm} \cos(x_n + x_n + x_n)]$$

If the calculated SE is less than 1 then increasing returns to scale, implying economies of scale. On the other hand, if SE = 1 then constant returns to scale and if SE is greater than 1 then decreasing returns to scale, implying diseconomies of scale.

Calculation of Scale inefficiency

Evanoff and Israilevich (1995) have noted that comparing scale economies (scale elasticities) with x-inefficiencies are misleading as the former is an elasticity and the latter is a relative efficiency measure. While many authors compare scale economies and x-inefficiencies, Evanoff and Israilevich suggest one should calculate scale inefficiencies for accurate comparisons.

The scale elasticity measure, $\varepsilon = \partial \ln C / \partial \ln Y$, is an elasticity associated with a particular output level and indicates the relative change in cost associated with an increment change from this output level. Scale inefficiency (I), on other hand, can be measured as the aggregate cost of N inefficient firms ($\varepsilon \neq 1.0$) relative to the cost of a single efficient firm ($\varepsilon = 1.0$); that is $I = [N * C_I / C_E] - 1.0$, where C_I and C_E are the cost of production at the inefficient and efficient firms, respectively.

Therefore, the two concepts differ because elasticity is related to incremental changes in output, and inefficiency related to the change in output required to produce at the minimum efficient scale. The inefficiency measure is typically associated with significantly larger output changes as it measures the difference in total or average cost at distinct output levels. Furthermore, the cost savings realised by an incremental increase in output by a scale inefficient firm is irrelevant for measuring inefficiency since this is not the savings realised by producing at the efficient scale.

Given the following simple representation for the cost function:

$$\ln C = a + b(\ln Y) + .5c(\ln Y)^2,$$

then the scale elasticity for inefficient firms $= \varepsilon_I = \partial \ln C_I / \partial \ln Y_I = b$, on the other hand the scale elasticity for the efficient firms = 1.0; by definition.

The scale inefficiency (see Evanoff and Israilevich, 1995) then can be written as:

$$I = e^{(.5/c)(1-\varepsilon_I)^2} - 1.0,$$

that is scale inefficiency is a function of the first and second derivatives of the function (cost function as well as other functional forms) with respect to output (the second derivation aims to reach c which is the key for inefficiency calculation).

Furthermore, if the estimated scale elasticity value is insignificantly different from unity, this does not imply scale inefficiency is insignificantly different from zero because the statistical difference of the elasticity measure from a value of unity depends entirely on the standard error of the estimated coefficient b .

For completeness, this study estimates x-inefficiencies, scale inefficiencies and scale economies for our sample of Arabic banks.

3. Data

Our data comprise a representative sample of the banks operating in Jordan, Egypt, Saudi Arabia and Bahrain and consist of 82 banks over the 1992-2000 period. This sample represents around 78%, 88%, 63% and 55% of the financial systems of these countries (excluding the assets of foreign branches and central banks) (Table 1 below shows the details).

Table 1

Size of the study sample relative to the banking sectors of Jordan, Egypt, Saudi Arabia and Bahrain over 1992-2000 (US\$ million, figures rounded to nearest 2 digits)

Country/ Year	Bahrain			Egypt			Jordan			Saudi Arabia		
	Sample Assets	Total Banking Assets	%	Sample Assets	Total Banking Assets	%	Sample Assets	Total Banking Assets	%	Sample Assets	Total Bank- ing Assets	%
1992	34,200	77,500	44	52,200	62,500	84	6,900	9,100	75	77,600	129,600	60
1993	34,300	68,400	50	54,300	60,900	89	7,100	9,600	74	82,700	142,800	58
1994	37,000	73,700	50	57,200	62,300	92	8,000	10,700	75	85,400	146,300	58
1995	40,000	73,700	54	63,900	69,800	92	9,100	11,900	77	89,600	150,100	60
1996	42,500	76,600	55	67,600	77,100	88	9,800	12,500	79	93,900	156,400	60
1997	44,900	83,500	54	77,200	89,100	87	11,100	13,700	81	105,000	163,900	64
1998	48,700	99,400	49	82,600	97,300	85	12,000	14,800	81	111,500	171,400	65
1999	55,200	102,100	54	88,700	103,300	86	13,000	16,300	80	121,700	172,200	71
2000	57,400	106,400	54	93,800	103,600	90	14,500	18,900	77	131,900	181,300	73
Average	43,800	84,600	52	70,800	80,600	88	10,200	13,100	78	99,900	157,100	63

Source: The total assets were extracted from the annual financial reports of the monetary agencies in the countries under study (the consolidated financial statements of the banks) while the sample was drawn from the London Bankscope database (January, 2000 & 2002).

Our sample represents the major financial institutions that have consistently published their financial statements over the last ten years in the countries under study. The relative size of Bahrain's banks sample looks small and the reason is that the financial system in this country has been dominated by offshore banking units which are excluded from the sample as these belong to large international financial institutions and their data are unavailable. In Saudi Arabia, the specialised

government institutions, while important, do not publish detailed financial statements and so these are not included in the sample.

Table 2 shows the specialisation of the banks included in the sample. The number of commercial banks comprises around 66% of the total sample. The percent of commercial banks operating in each country varies ranging from 42% in Bahrain to 77% in Saudi Arabia.

Table 2

Specialisation of banks under study, 1992-2000

% of total	Bahrain	Egypt	Jordan	Saudi Arabia	All
Commercial	44	76	57	77	66
Investment	28	8	29	8	16
Islamic	17	5	7	0	7
Other	11	11	7	15	11
Total Number	18	37	14	13	82

Source: Bankscope (Jan. 2000 & 2002).

Table 3 shows that the size of total assets of all the banks included in the present study increased from about US\$ 180 billion in 1992 to about US\$ 310 billion in 2000 and averaged about US\$ 235 billion over the whole period. Dividing these financial institutions into nine size categories, the share of the largest banks (with assets size greater than US\$ 5 billion) constituted around 70 percent of the total assets of all the banks over the period of 1992-2000.

Table 3

Distribution of banks' assets in Jordan, Egypt, Saudi Arabia and Bahrain, 1992-2000

	1992	1993	1994	1995	1996	1997	1998	1999	2000	Avg.
	%	%	%	%	%	%	%	%	%	US\$, mil.
1-99.9	0.11	0.08	0.14	0.16	0.14	0.10	0.06	0.02	0.02	202
100-199.9	1.16	1.05	0.78	0.35	0.31	0.18	0.21	0.29	0.27	1,073
200-299.9	1.76	1.35	1.10	1.78	1.04	0.80	0.67	0.36	0.32	2,173
300-499.9	3.78	4.08	3.47	2.79	2.92	2.75	2.49	2.04	1.58	6,422
500-999.9	2.56	2.73	4.64	4.57	4.51	3.53	3.67	3.47	3.29	8,569
1,000-2,499.9	11.87	11.50	9.89	13.09	10.02	11.31	11.84	10.51	10.15	25,911
2,500-4,999.9	8.29	8.56	4.68	4.94	7.12	6.65	6.50	7.66	8.26	16,470
5,000-9,999	18.22	19.28	24.51	26.23	24.40	26.82	14.88	19.13	9.28	46,196
10,000+	52.26	51.37	50.78	54.22	49.54	47.85	59.67	56.53	66.83	129,190
T. Assets (US\$, mil., nominal values)	179,033	186,975	197,046	213,044	225,426	250,325	267,943	292,855	313,209	

Source: Bankscope (Jan. 2000 & 2002).

This study employs the intermediation approach for defining bank inputs and outputs. Following Aly et al. (1990), the inputs used in the calculation of the various efficiency measures are deposits (W_1), labour (W_2) and physical capital (w_3). The deposits include time and savings deposits, notes and debentures, and other borrowed funds. The price of loanable funds was derived by taking the sum of interest expenses of the time deposits and other loanable funds divided by loanable

funds. Labour is measured by personnel expenses as a percent of total assets¹. Bank physical capital is measured by the book value of premises and fixed assets (including capitalised leases). The price of capital was derived by taking total expenditures on premises and fixed assets divided by total assets. The three outputs used in the study include total customer loans (y_1), all other earning assets (y_2), and off-balance sheet items (y_3), measured in millions of US dollars.

The off-balance sheet items (measured in nominal terms) were included as a third output. Although the latter are technically not earning assets, these constitute an increasing source of income for banks and therefore should be included when modelling the banks' cost characteristics; otherwise, total banks' output would tend to be understated (Jagtiani and Khanthavit, 1996). Furthermore, these items are included in the model because they are often effective substitutes for directly issued loans, requiring similar information-gathering costs of origination and ongoing monitoring and control of the counterparts, and presumably similar revenues as these items are competitive substitutes for direct loans.

The definitions, means, standards of deviation of the input and output variables used in the stochastic frontier estimations are reported in Table 4. The table shows that the average bank had US\$ 1.26 billion in loans, US\$ 1.39 billion other earning assets and US\$ 1.32 billion of balance sheet items over 1992-2000. The cost of input variables averaged about 7.0 percent for purchased funds, 2.0 percent for labour and 1.0 percent for physical capital over the period of 1992-2000. On the other hand, the prices of banks output averaged about 15.0 percent for loans²; 5.0 percent for other earning assets and 1.0 percent for off-balance sheet items over the same period.

Table 4

Descriptive statistics of the banks' inputs and outputs for Jordan, Egypt, Saudi Arabia and Bahrain over 1992-2000

Variables	Description	Mean	St. Dev.	Min.	Max.
TC	Total cost (includes Interest expense, Personnel expense, Commission expense, Fee expense, Trading expense, other operating expense) (US\$ millions).	170	300	0	1,720
W1	Price of funds (%) (total interest expense/ total customer deposits (demand, saving and time deposits)).	0.07	0.09	0.00	1.98
W2	Price of labour (%) (total personnel expense/total assets).	0.02	0.01	0.00	0.21
W3	Price of physical capital (Non-interest expense/Average assets).	0.01	0.01	0.00	0.21
Y1	The US \$ value of total aggregate loans (all types of loans) (US\$ millions).	1,260	2,280	1	15,060
Y2	The US \$ value of total aggregate other earning assets (short-term investment, equity and other investment and public sector securities (US\$ millions)).	1,390	2,470	1	13,600
Y3	The US \$ value of the off-balance sheet activities (nominal values, US\$ millions).	1,320	3,510	1	26,740
p1	Price of loans (%) (total earned interest/ Total loans).	0.15	0.07	0.01	0.87
p2	Price of other earning assets (%) (Trading income and other operating income excluding commission and fees income/Other earning assets).	0.05	0.04	0.01	0.33
P3	Price of off-balance sheet items (%) (Commission and fees income/ off-balance sheet items).	0.01	0.02	0.00	0.20

Source: Bankscope (Jan. 2000 & 2002).

¹ As staff numbers were not available for the banks in the sample, we used this measure instead. This measure for staff costs has been used in various previous studies including Altunbas et al. (1996) and (1999).

² This may be an overstatement as interest earned on bonds is also included in this figure.

In addition to the above input and output variables, the present study employs a variety of control and environmental variables¹ to rule out the effect of other factors that might explain differences among efficiency estimates for the banks under study. The three control variables included in our model include the size of loan loss reserves as a percent of bank's credit portfolio, the capital adequacy ratio, and a time trend (see Table 5 below for details). The loan loss reserves as a proportion of gross loans ranged between 0.01 and 19.68 percent, the latter figure suggests that some banks faced substantial credit quality problems. The total banks' capital as a percentage of total assets averaged around 14.0 percent with a standard deviation of 12.0 percent, this reflects sizeable differences in the capital adequacy of the banks under study.

The size of loan loss reserves as a proportion of gross loans is added to the model to control for the bank's risk structure. It is also used as a measure of bank's asset quality and as a measure of the bank's management efficiency in monitoring the credit portfolio. A lack of diversity in a bank's asset portfolio may be associated with increases in problem loans without sufficient provisioning, exposing bank's capital to risk and potential bankruptcy that might be closely related to the quality of bank management. Banks facing financial distress have been found to carry large proportions of nonperforming loans (Whalen, 1991). Furthermore, studies on bank failures suggest a positive relationship between operating inefficiency and failure rates (see for example, Cebenoyan, Cooperman, and Register, 1993; Hermalin and Wallace, 1994; Wheelock and Wilson, 1995). Barr, Seiford and Siems (1994) found that this positive relationship between inefficiency and failure is evident a number of years ahead of eventual failure. Kwan and Eisenbeis (1994) report that problem loans are negatively related to efficiency even in non-failing banks. Berger and DeYoung (1997) found a link between management quality and problem loans by reporting that an increase in management quality reduces the bank's problem loans.

Hughes et al. (1996a, b) and Mester (1996) included the volume of nonperforming loans as a control for loan quality in studies of US banks, and Berg et al. (1992) included loan losses as an indicator of loan quality evaluations in a DEA study of Norwegian bank productivity. Whether it is appropriate to include nonperforming loans and loan losses in bank's cost, standard and alternative profit functions depends on the extent to which these variables are exogenous. Such variables would be exogenous if caused by negative economic shocks "bad luck", but they could be endogenous, either because management is inefficient in managing its portfolio "bad management" or because it has made a conscious decision to reduce short-run expenses by cutting back on loan origination and monitoring resources "skimping". Berger and DeYoung (1997) tested the bad luck, bad management, and skimping hypotheses and found mixed evidence on the exogeneity of nonperforming loans.

Another important aspect of efficiency measurement is the treatment of financial capital. A bank's insolvency risk depends on the financial capital available to absorb portfolio losses, as well as on the portfolio risk themselves. Even apart from risk, a bank's capital level directly affects costs by providing an alternative to deposits as a funding source for loans. On the other hand, raising equity typically involves higher costs than raising deposits. If the first effect dominates, measured costs will be higher for banks using a higher proportion of debt financing; if the second effect dominates, measured costs will be lower for these banks. The specification of capital in the cost and profit functions also goes part of the way toward accounting for different risk preferences on the parts of banks. Therefore, if some banks are more risk averse than others, they may hold a higher level of financial capital than maximising profits or minimising costs. If financial capital is ignored, the efficiency of these banks would be mismeasured, even though they behave optimally given their risk preferences. Hughes et al. (1996a, b, 1997) and Hughes and Moon (1995) tested and rejected the assumption of risk neutrality for banks. Clark (1996) included capital in a model of economic cost and found that it eliminated measured scale diseconomies in production costs

¹ The control variables enter into the stochastic frontier model in the same way as the input variables (as betas) and these variables are fully interactive with other parameters of the model; On the other hand, the environmental variables are not interactive with other model parameters and added to the model as delta (as will be shown later).

alone. The cost studies of Hughes and Mester (1993) and the Hughes et al. (1996a, 1997) profit studies incorporated financial capital and found increasing returns to scale at large-asset-size banks. A possible reason is that large size confers diversification benefits that allow large banks to have lower capital ratios than smaller banks. Akhavein et al. (1997a) controlled for equity capital and found that profit efficiency increases as a result of mergers of large banks. Bank's capital is also included in the model of Berger and Mester (1997) who find that well-capitalised firms are more efficient. This positive relationship between capital and efficiency may indicate that inefficient banks with lower capital have less to lose in taking more risky projects than an efficient bank. This is consistent with moral hazard and agency conflict between managers and shareholders where less monitored managers with lower equity have incentives to expense preference.

The environmental variables (or efficiency correlates) were also added to the model to investigate the reason for the differences in efficiency scores across banks under study. These include variables that control for market structure and organisational characteristics, geographical segmentation and bank liquidity. We identify variables to account for bank specialisation, bank size and concentration in the respective banking industries. Financial institutions in each country are divided into four categories; commercial, investment, Islamic and other financial institutions (that perform various bank functions). Furthermore, we employ the 3-firm asset concentration ratio which is widely used to test for monopoly characteristics. Furthermore, we include a dummy variable to control for bank geographical (countries) location (Table 5 shows descriptive statistics of the control and environmental variables).

Table 5

Descriptive statistics of the banks' control and environmental variables for Jordan, Egypt, Saudi Arabia and Bahrain over 1992-2000

Variables	Description	Mean	St. Dev	Min	Max
<u>The control Variables</u>					
K	Capital Adequacy (%) (Total equity/Total Assets)	0.14	0.12	0.01	0.72
S	Asset quality (Loan Loss Reserve/Gross Loans)	0.22	0.81	0.01	19.68
T	Time Trend	5.00	2.58	1.00	9.00
<u>The Environmental Variables</u>					
TA	Total Assets (US\$ millions)	2,881	4,966	35	26,700
B	Dummy variable for Bahrain	0.22	0.41	0.00	1.00
J	Dummy variable for Jordan	0.17	0.38	0.00	1.00
E	Dummy variable for Egypt	0.45	0.50	0.00	1.00
Com.	Dummy variable for commercial banks	0.66	0.47	0.00	1.00
Inv.	Dummy variable for investment/securities banks	0.16	0.37	0.00	1.00
Isl.	Dummy variable for Islamic banks	0.07	0.26	0.00	1.00
L	Liquidity ratio (%) (Total liquid assets/Total Assets)	0.14	0.16	0.00	0.71
3-FCR	Three firm concentration ratio (%) (the largest 3 banks total assets of/Total assets of all banks in the bank country for the respective years)	0.62	0.14	0.48	0.81
MS	Bank assets market share (%) for each year	0.05	0.10	0.00	0.68

Source: Bankscope (Jan., 2000, 2002).

The total assets variable is used to control for bank size where bank size should be strongly associated with efficiency as size may be required to utilise scale and (maybe) scope economies (if large banks are more diversified). Furthermore, larger banks may have more professional management teams and/or might be more cost conscious due to greater pressure from owners concerning the bottom line profits (Evanoff and Israilevich, 1991). Berger et al. (1993) found that most of the efficiency differences among large banks was on the output side as larger banks might be better able to reach

their optimal mix and scale of outputs. On the other hand, Hermalin and Wallace (1994), Kaparakis et al. (1994), DeYoung and Nolle (1996) found significant negative relationships. Other studies, however, report no significant relationship between bank size and efficiency, such as Aly et al. (1992), Cebenoyan et al. (1993), Mester (1993), Pi and Timme (1993), Mester (1996), Berger and Hannan (1995), Berger and Mester (1997), and Chang et al. (1998).

The 3-firm concentration ratio and market share variables were included to control for oligopoly behaviour along the lines of the traditional structure-conduct-performance paradigm (see Molyneux et al., 1996) and as an indicator of the characteristics of the respective banking industry structures. The Cournot model of oligopolistic behaviour suggests that there is a positive relationship between concentration and profitability. Consistent with this model, some studies have found a positive relationship between market concentration and profitability (Berger and Hannan, 1997; Berger and Mester, 1997). The market power that prevails in the less competitive markets enables some banks to charge higher prices for their services and make supernormal profits. Banks may exert their own market power through size as noted by Berger (1995) and so we include a market share variable to control for what Berger refers to as 'relative market power'.

Dummy variables for bank specialisation are also included in the model so as to control for the product diversity as efficiency might associated with firm's strength in carefully targeting its market niches. The cost of producing various products might be lower when specialised banks produce them rather than when a single bank produces all the products due to diseconomies of scope. There are number of studies that have examined the impact of product diversity on efficiency. Aly et al. (1990) found a negative relationship between product diversity and cost efficiency. Ferrier, Grosskopf, Hayes and Yaisawarng (1993) found that banks with greater product diversity tend to have lower cost efficiency. Chaffai and Dietsch (1995) compared the efficiency of universal versus non-universal (more specialised) banks in Europe and found the former to be less cost efficient.

Finally, the liquidity ratio is included to account for bank's liquidity risk. Banks that hold more liquidity may be expected to have lower liquidity risk but may be less profit efficient as liquid assets tend to yield lower returns. In contrast, as liquid assets are controlled in outputs, one would expect banks with higher liquid assets (all other things being equal) to be more cost efficient.

4. Results: Efficiencies and Productivity Changes

This section presents the steps undertaken our preferred cost model. This includes employing different models utilised in the banking efficiency literature based on different assumptions concerning the distribution of efficiency terms. In addition, various hypotheses are tested, given different combination of control and environmental variables, to arrive at the preferred models based on maximum likelihood estimation¹. Based on the preferred model, we present cost efficiency, scale elasticity and scale efficiency measures for the banks under study.

¹ The Maximum Likelihood (ML) and Log-likelihood (LL) functions are the basis for deriving parameters estimates, given certain data. While the shapes of these two functions are different, they have their maximum point at the same value. Both seek to estimate the value of p (the unknown parameter in the model) that maximises the ML or LL function given the data z . The MLEs have many statistical appealing features especially when the sample size is large. First, consistency: as the sample size increases, the MLEs converge to the true parameters values. Second, asymptotic normality and efficiency (i.e., as the sample size increases, the sampling distribution of the MLE converges to normality with least possible variance (Hence, estimates obtained typically have the smallest confidence intervals)). The MLE of unknown parameter, \hat{p} is the value of p that corresponds to the maximum of $L(p/z)$ that is most likely to have produced from data z . Since it is easier to deal with addition rather than multiplication, the problem is generally tackled in the log form. This is called the log likelihood function that truly maximises the sum of the log likelihoods by choosing the parameters that give identical results to maximising the untransformed likelihood. The log likelihood takes the following form:

$$\log(L) = -(n/2)\log(2\pi) - (n/2)\log(\sigma_u^2) - (1/2)\sigma_u^2 \sum_i (Y_i - a - bX_i)^2.$$

There are three stages undertaken to arrive at the preferred model for our cost function estimates. The first stage involves utilising Battese and Coelli's (1995) approach that allows us to include the efficiency correlates directly in the model estimation. The second stage involves utilising Battese and Coelli's (1992) time-varying efficiency approach that gives flexibility to examine different assumptions concerning the distribution of efficiency terms, comparing time-variant versus time-invariant models but it does not allow for the inclusion of efficiency correlates in the model. Finally, stage 3 compares the best specified models in stage 1 and stage 2 to arrive at a single preferred model from the two stages and provides the basis for the model choice.

Stage 1: Estimating the cost frontier models that include efficiency correlates

This stage estimates the stochastic frontier for the cost function, given the Fourier-flexible functional form that includes efficiency correlates. This stage follows Dietsch and Lozano-Vivas (2000) who emphasise the importance of including country and other specific information in common frontier estimations of bank efficiency. This stage is conducted using the approach suggested by Battese and Coelli's (1995) technical inefficiency effects model that allows us to include firm-specific (and country-specific variables) directly into the model as these might explain some of the efficiency differences between banks as well as the variation in bank inefficiency overtime.

Battese and Coelli's (1995) model defines the inefficiency term u_{it} as non-negative variables that account for technical inefficiency and are assumed to be independently and identically distributed (iid) as truncations at zero of the $N(\delta_{it}d, \sigma_{u_{it}}^2)$ distribution. This methodology follows Kumbhakar, Ghosh and McGukin (1991) and Reifschneider and Stevenson (1991) and Battese and Coelli (1991) who propose a stochastic model in which u_{it} are stated as an explicit function of a vector of firm-specific variables and random error. According to Coelli (1996), this specification proves to be better than that of Pitt and Lee (1981) who have estimated stochastic frontiers and predicted firm-level efficiencies using these estimated functions, and then regressed the predicted efficiencies upon firm-specific variables (such as managerial experience, ownership characteristics, etc.) in an attempt to identify some of the reasons for differences in predicted efficiencies between firms. Furthermore, the two-stage procedure utilised by Pitt and Lee (1981) has been recognised as one which is inconsistent in its assumptions regarding the independence of the inefficiency effects in the two estimation stages.

In order to derive the bank efficiency model that includes firm-specific variables, we employ the control and environmental variables detailed earlier. The control variables include the loan loss reserves as a percent of loans, capital strength and a time trend. The loan loss reserve as a percent of gross loans is included to control for asset quality. Capital strength is measured by the ratio of equity to total asset ratio. A time trend variable is included in the model (Table 5 shows descriptive statistics of these variables). Environmental variables are employed, as a set of explanatory variables, to control for organisational characteristics, geographical location. Organisational characteristics refer to the structure of the financial systems in the countries under study. We identify three ratios to test these characteristics: dummy variables for bank specialisation, bank market share and concentration in the pertinent banking systems. The banks in each country are divided into four categories; commercial, investment, Islamic and other financial institutions. Furthermore, we employ the 3-firm concentration ratio which is widely used to test for monopoly characteristics in the pertinent market. Furthermore, we include dummy variables to control for bank geographical (country) location.

To reach the best-specified model in this stage, we have examined many hypotheses which can be summarised in the following steps:

Step 1: Estimating the Fourier-truncated with different combination of control variables (see Table 6 for details)

1.1. The unrestricted Fourier-flexible model is estimated assuming inefficiency to be truncated. This model includes all the control variables (bank's capital, bank's asset quality and the time

trend) and all the efficiency correlates (the environmental variables). This general model will be compared later with some other models to decide upon (based on maximum-likelihood ratio tests) preferred model specifications utilising different combinations of control variables.

I.2. The Fourier-truncated model that includes the efficiency correlates is estimated but without the time parameters. This is done to examine whether there has been any technical change over the sample period. This involves restricting all the coefficients associated with the time trend equal to zero. Next, we estimate the model but without the capital parameters. Then, we estimated the model without the risk (bank's asset quality) parameters.

At this point, there are three null hypotheses to be examined. The first null hypothesis is that the specification of the truncated model without time parameters is better than that of the unrestricted model in (I.1). The second null hypothesis states that the specification of the truncated model without the risk parameters is better than that of the unrestricted model. The third null hypothesis states that specification of the truncated model without capital parameters is better than that of the unrestricted model. The alternative hypothesis (H_a) against these hypotheses is that the full model (I.1) is better specified than these restricted models.

As Table 6 shows, based on the log-likelihood one-sided ratio¹, only the null hypothesis that the model without time-parameters is better specified model is accepted at the critical value of 5% while the other null hypotheses are rejected. In other words, the value of the generalised likelihood-ratio statistics compared with those of the upper five per cent point for χ -square (for the appropriate degree of freedom) were not in favour of accepting these null hypotheses. This means that the model without time parameters is better specified than the unrestricted model (I.1 above).

I.3. The Fourier-truncated model that includes the efficiency correlates is estimated without time and capital parameters simultaneously. Next, the model is estimated without time and risk parameters. Then, the model is estimated without risk and capital parameters.

Again here, we have three null hypotheses that need to be examined. The first null hypothesis states that the Fourier-truncated that includes the efficiency correlates but without time and capital parameters is specified better than the models in I.1 and I.2 above. The second null hypothesis states that the truncated model without time and risk parameters is better specified than those in I.1 and I.2. Finally, the third null hypothesis states that the truncated model without risk and capital parameters is better specified than those in I.1 and I.2. Based on the log-likelihood ratio, all the null hypotheses are rejected (Table 6 shows the details).

I.4. The Fourier-truncated that includes the efficiency correlates is estimated but without any of the control variables (capital, risk and time) in the model. In this case, the null hypothesis states that Fourier-truncated model excluding the control variables is specified better than the models specified in I.1, I.2 and I.3 above. Based on the maximum likelihood ratio, this model is not rejected at critical level of 5%. Therefore, the best specified model up to this step is the Fourier-truncated that excludes all the control variables.

¹ The Maximum likelihood (ML) provides a convenient way to test the hypotheses in the form of the Log-likelihood ratio (LR) that examine whether a reduced model provides the same fit as a full model. This ratio allows us to test whether the likelihood estimates for parameters are significantly different from other fixed values. It permits to compare the likelihood of the data under one hypothesis against the likelihood of the data under another (more restricted) hypothesis. The LR shows whether the data are significantly less likely to have arisen if the null hypothesis is true than if the alternate hypothesis is true? The difference between the likelihoods is multiplied by a factor of 2 for technical reasons, so that this quantity will be distributed as the familiar χ^2 statistic. The LR test statistic is given by $LR = -2[L(\hat{\theta}_r / z) - L(\hat{\theta} / z)]$ where $L(\hat{\theta} / z)$ is the likelihood function evaluated at the MLE where $L(\hat{\theta}_r / z)$ is the maximum if the likelihood function, subject to the restriction that r unconstrained parameters in the full likelihood analysis are assigned fixed values. For sufficiently large sample size, the LR test statistic is χ_r^2 -distributed, a χ^2 with r degrees of freedom (Wald, 1943). The degrees of freedom equal the difference in the number of parameters being estimated under the alternate and null models.

Table 6

Hypotheses testing of the cost function (stage 1)

Model Description	Restrictions	Log likelihood	LR test of 1-sided error	DF	Critical value for $\alpha = 5\%$	Decision
Stage 1: Models estimation including environmental variables						
- Fourier-truncated without restrictions		108.02				
- Fourier-truncated without time parameters	$\Psi_3 = \Psi_{r3} = \Psi_{3S} = \rho_{i3} = \tau_{k3} = \phi_8 = \omega_8 = \phi_{n8} = \phi_{8q} = \omega_{8q} = \omega_{n8} = \phi_{888} = \omega_{888} = 0, i=S=k=1, 2, 3; j=1, 2; n=q=1, 2, \dots, 8.$	193.42	-170.8	29	42.56	Accept Ho
- Fourier-truncated without capital parameters	$\Psi_1 = \Psi_{r1} = \Psi_{iS} = \rho_{i1} = \tau_{k1} = \phi_6 = \omega_6 = \phi_{n6} = \phi_{6q} = \omega_{6q} = \omega_{n6} = \phi_{666} = \omega_{666} = 0, i=S=k=1, 2, 3; j=1, 2; n=q=1, 2, \dots, 8.$	13.29	189.46	29	42.56	Reject Ho
- Fourier-truncated without risk parameters	$\Psi_2 = \Psi_{r2} = \Psi_{2S} = \rho_{i2} = \tau_{k2} = \phi_7 = \omega_7 = \phi_{n7} = \phi_{7q} = \omega_{7q} = \omega_{n7} = \phi_{777} = \omega_{777} = 0, i=S=k=1, 2, 3; j=1, 2; n=q=1, 2, \dots, 8.$	69.07	77.9	29	42.56	Reject Ho
- Fourier-truncated without time and capital parameters	$\Psi_1 = \Psi_3 = \Psi_{r1} = \Psi_{r3} = \Psi_{iS} = \Psi_{3S} = \rho_{i1} = \rho_{i3} = \tau_{k1} = \tau_{k3} = \phi_6 = \phi_8 = \omega_6 = \omega_8 = \phi_{n6} = \phi_{n8} = \phi_{6q} = \phi_{8q} = \omega_{6q} = \omega_{8q} = \omega_{n6} = \omega_{n8} = \phi_{666} = \phi_{888} = \omega_{666} = \omega_{888} = 0, i=S=k=1, 2, 3; j=1, 2; n=q=1, 2, \dots, 8.$	-80.17	547.18	26	38.88	Reject Ho
- Fourier-truncated without time and risk parameters	$\Psi_2 = \Psi_3 = \Psi_{r2} = \Psi_{r3} = \Psi_{2S} = \Psi_{2S} = \Psi_{iS} = \rho_{i2} = \rho_{i3} = \tau_{k2} = \tau_{k3} = \phi_7 = \phi_8 = \omega_7 = \omega_8 = \phi_{n7} = \phi_{n8} = \phi_{7q} = \phi_{8q} = \omega_{7q} = \omega_{8q} = \omega_{n7} = \omega_{n8} = \phi_{777} = \phi_{888} = \omega_{777} = \omega_{888} = 0, i=S=k=1, 2, 3; j=1, 2; n=q=1, 2, \dots, 8.$	161.26	64.32	26	38.88	Reject Ho

Step 2: Comparing Fourier specification with translog specification

In this step, we will compare the best Fourier specifications concluded from **step 1** with identical translog specifications. The null hypothesis in this step states that translog specifications are more appropriate than the Fourier specifications for estimating efficiency. The alternative hypothesis states that translog specification is not better than that of the Fourier. Based on the log-likelihood ratio, the null hypothesis is rejected at the 5% significance level. This means that the data are better specified utilising the Fourier than the translog form.

Step 3: Examining the impact of efficiency correlates (the environmental variables) on the model specification

The best specified model up to **step 1** and **2** above is the Fourier-truncated that includes the efficiency correlates (environmental variables) but does not include any of the control variables. In the following, we estimate the Fourier-truncated without including the efficiency correlates. In this case, the null hypothesis states that the specified truncated model without efficiency correlates is better than the model that includes them. The alternative hypothesis, on the other hand, states that the model that excludes the efficiency correlates is not specified better than the model that includes them. Based on the log-likelihood ratio, the null hypothesis is rejected in favour of the alternative hypothesis that necessitates the existence of such variables in the model (see Table 6 for details).

Step 4: Examining the impact of inefficiency-terms on the model specification

In this step, the best specified model selected until **step 3** will be compared with the model that excludes the inefficiency term from the model. The null hypothesis here states that the inefficiency effects in the cost function are not present, and so the banks are fully technically efficient. If this is the case, the technical inefficiency error term, U_{it} , would be removed from equation, and the resulting model would be appropriately estimated using OLS. This hypothesis is rejected and so, the model which accounts for technical inefficiency is warranted in these instances (see Table 6 for details).

Based on the results of the steps above, the best specified model from **stage 1** is the Fourier-truncated model that excludes the control variables (time trend, capital adequacy and asset quality) but includes the efficiency correlates (Table 6 shows the details).

Stage 2: Estimating the cost frontier models that excludes efficiency correlates

This stage estimates the stochastic frontier, given the Fourier-flexible functional form that excludes efficiency correlates. The models in this stage are estimated utilising Battese and Coelli's (1992) time-varying approach. This approach gives some flexibility concerning the distribution of inefficiency term in the stochastic frontier; truncated or half normal. Furthermore, it allows us to examine the time-varying efficiency model against the time-invariant model. Therefore, one of the advantages of the time-varying inefficiency model is that the technical inefficiency changes over-time can be distinguished from technical change, provided the latter is specified in the model parameters, in the frontier function. This discrimination is only possible given that the technical inefficiency effects are stochastic and have the specified distributions. However, this approach does not allow us to add the efficiency correlates directly into the model.

The inefficiency term u_{it} s in this model is assumed to be an exponential function of time, involving only one unknown parameter. The technical inefficiency effects are assumed to be defined by

$$u_{it} = \{exp[-\eta(t - T)]\}u_i, \quad i = 1, 2, \dots, N; t = 1, 2, \dots, T;$$

where u_{it} s are assumed to be independently and identically distributed (i.i.d.) as the generalised truncated-normal random variable and η is an unknown scalar parameter to be estimated. The major disadvantage of this time-varying model is that the technical inefficiency effects of different

firms at any given time period, t , are equal to the identical exponential function ($\exp[-\eta(t-T)] \equiv \exp[\eta(T-t)]$) of the corresponding firm-specific inefficiency effects at the last period of the panel (the u_{it} s). This implies that the ordering of the firms according to the magnitude of the technical inefficiency effects is the same at all time periods. Thus, the time-varying model of the equation does not account for situations in which some firms may be relatively inefficient initially but become relatively more efficient in subsequent periods.

In our search for the best model specification utilising this model, we follow studies that assume no restriction to be imposed on the distributional features of the inefficiency term. These studies include Cebenoyan et al. (1993) who use the truncated normal model, Stevenson (1980) and Greene (1990) who use the normal and gamma distribution respectively. Then, we restricted μ (μ) to be zero to obtain Pitt and Lee's (1981) half-normal model. The studies that use the half-normal specification to model inefficiency in banking include Allen and Rai (1996), Kaparakis et al. (1994) and Mester (1996). Next, we restrict both μ (μ) and η (η) to be zero to get the time-invariant model as outlined in Battese, Coelli and Colby (1989). All the above models assume that the inefficiency term to be independently and identically as truncations at zero of the $N(\mu, \sigma_u^2)$ distribution. This definition of the inefficiency term conforms to the original definition of the stochastic frontier, which was proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and Van de Broeck (1977).

The following steps summarise the procedures followed to arrive at the most appropriate model specifications in this stage using Battese and Coelli's (1992) approach:

Step 1: Comparing the Fourier-truncated time-variant with time-invariant model

The specification of the estimated truncated time-variant model is compared with the truncated time-invariant model and the better specified model is chosen based on the log-likelihood ratio test. The null hypothesis in this step states that the specification of Fourier-truncated time-invariant model is better than the time-variant model. The null hypothesis in this step is rejected, as the time-invariant model cannot be specified using the stochastic frontier methodology (see table 7 for details).

Step 2: Fourier truncated time-variant versus Fourier half-normal time variant model

The specification of the truncated time-variant model chosen from step 1 is compared with the half-normal time-variant model. Here, the null hypothesis states that the half-normal time-variant model specification is better than the specification of the truncated time-variant model. Utilising the log-likelihood ratio, the null hypothesis is rejected given the appropriate degree of freedom.

Step 3: Fourier-truncated with different combinations of control variables

The Fourier-truncated time-variant model is estimated with different combinations of control variables to see if we can accept simpler model specification for our data. In this step, there are seven hypotheses examined. The first one states that the specification of the Fourier-truncated time-variant model without time parameters is better than the model specified in *steps 1* and *2* above. The second hypothesis examines the model without risk parameters and the third one examines the model without the capital parameters. The fourth hypothesis examines the model without time and risk parameters at the same time. The fifth hypothesis examines the model without time and capital parameters. The sixth hypothesis examines the model without capital and risk parameters. Finally, the seventh hypothesis examines the model specification without any of the control variables (capital, risk and time trend). Comparing the estimated models in this step and based on the log-likelihood ratio, the most appropriate model is the Fourier-truncated time-variant model without the control variables (see Table 7 for details).

Table 7

Hypotheses testing of the cost function (stage 2)

Model Description	Restrictions	Log likelihood	LR test of 1-sided error	DF	Critical value for $\alpha = 5\%$	Decision
Stage 2: Models estimation excluding environmental variables						
Step 1: Time-variant versus time-invariant models						
- Truncated time-variant model that includes all the control variables		114.42				
- Truncated time-invariant model that includes all the control variables	$\eta = 0$	ols				Reject Ho
Step 2: Truncated versus half-normal models						
-Half-normal time-variant model that includes all the control variables	$\mu = 0$	111.19	7.45	1	3.841	Reject Ho
Step 3: Truncated time-variant model with different combination of the control variables						
- Fourier-truncated without time parameters	$\Psi_3 = \Psi_{r,3} = \Psi_{3S} = \rho_{13} = \tau_{k,3} = \phi_8 = \omega_8 = \phi_{n,8} = \phi_{8q} = \omega_{n,8} = \omega_{n,8} = \phi_{888} = \omega_{888} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	-6.70	235.78	29	42.56	Reject Ho
- Fourier-truncated without capital parameters	$\Psi_1 = \Psi_{r,1} = \Psi_{1S} = \rho_{11} = \tau_{k,1} = \phi_6 = \omega_6 = \phi_{n,6} = \phi_{6q} = \omega_{n,6} = \omega_{n,6} = \phi_{666} = \omega_{666} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	29.10	170.64	29	42.56	Reject Ho
- Fourier-truncated without risk parameters	$\Psi_2 = \Psi_{r,2} = \Psi_{2S} = \rho_{12} = \tau_{k,2} = \phi_7 = \omega_7 = \phi_{n,7} = \phi_{7q} = \omega_{n,7} = \omega_{n,7} = \phi_{777} = \omega_{777} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	8.45	211.94	29	42.56	Reject Ho

Table 7 (continuous)

Model Description	Restrictions	Log likelihood	LR test of 1-sided error	DF	Critical value for $\alpha = 5\%$	Decision
- Fourier-truncated without time and capital parameters	$\Psi_1 = \Psi_3 = \Psi_{r1} = \Psi_{r3} = \Psi_{1S} = \Psi_{3S} = \rho_{11} = \rho_{13} = \tau_{k1} = \tau_{k3} = \phi_6 = \phi_8 = \omega_6 = \omega_8 = \phi_{n6} = \phi_{n8} = \phi_{n8} = \phi_{n8} = \phi_{6q} = \phi_{8q} = \omega_{6q} = \omega_{8q} = \omega_{n6} = \omega_{n8} = \phi_{666} = \phi_{888} = \omega_{666} = \omega_{888} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	-83.88	396.6	55	73.11	Reject Ho
- Fourier-truncated without time and risk parameters	$\Psi_2 = \Psi_3 = \Psi_{r2} = \Psi_{r3} = \Psi_{2S} = \Psi_{3S} = \rho_{12} = \rho_{13} = \tau_{k2} = \tau_{k3} = \phi_7 = \phi_8 = \omega_7 = \omega_8 = \phi_{n7} = \phi_{n8} = \phi_{n8} = \phi_{7q} = \phi_{8q} = \omega_{7q} = \omega_{8q} = \omega_{n7} = \omega_{n8} = \phi_{777} = \phi_{888} = \omega_{777} = \omega_{888} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	-17.11	263.06	55	73.11	Reject Ho
- Fourier-truncated without capital and risk parameters	$\Psi_1 = \Psi_3 = \Psi_{r1} = \Psi_{r2} = \Psi_{1S} = \Psi_{2S} = \rho_{11} = \rho_{12} = \tau_{k1} = \tau_{k2} = \phi_6 = \phi_7 = \omega_6 = \omega_7 = \phi_{n6} = \phi_{n7} = \phi_{n7} = \phi_{n7} = \phi_{6q} = \phi_{7q} = \omega_{6q} = \omega_{7q} = \omega_{n6} = \omega_{n7} = \phi_{666} = \phi_{777} = \omega_{666} = \omega_{777} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	3.44	221.96	55	73.11	Reject Ho
- Fourier-truncated without time, capital and risk parameters	$\Psi_r = \Psi_{rS} = \rho_{ir} = \tau_{kr} = \phi_6 = \phi_7 = \phi_8 = \omega_6 = \omega_7 = \omega_8 = \phi_{n6} = \phi_{n7} = \phi_{n8} = \phi_{6q} = \phi_{7q} = \phi_{8q} = \omega_{6q} = \omega_{7q} = \omega_{8q} = \omega_{n6} = \omega_{n7} = \omega_{n8} = \phi_{666} = \phi_{777} = \phi_{888} = \omega_{666} = \omega_{777} = \omega_{888} = 0, r=S=k=1, 2, 3; i=1, 2; n=q=1, 2, \dots, 8.$	69.06	90.72	78	99.62	Do not reject Ho
Step 3: Fourier-truncated versus translog						
- Translog-truncated without time, capital and risk parameters	$\Psi_r = \Psi_{rS} = \rho_{ir} = \tau_{kr} = \phi_n = \omega_n = \phi_{nq} = \omega_{nq} = \phi_{nmn} = \phi_{qqq} = 0, n=q=1, 2, \dots, 8.$	68.15	92.54	104	128.80	Accept Ho

Source: Author's own estimation.

Step 4: Comparing the Fourier-specification with translog specification

In this step, we compare the Fourier-truncated model specifications selected in **step 3** above with the translog form given an identical specification. At this point, the null hypothesis states that the translog specification is more appropriate than the Fourier specification. The null hypothesis is not rejected and so, the best specified model in this stage is the translog-truncated without the control variables.

Stage 3: Comparing the models from stage 1 and stage 2

It should be noted that we cannot formally compare directly the results of **stage 1** and **stage 2** above because we utilise Battese and Coelli's (1995) approach in the first stage and Battese and Coelli's (1992) approach in the second stage. The first approach does not have the second approach as a special case, and neither does the converse apply. Thus, these two model specifications are non-nested and hence no set of restrictions can be defined to permit a test of one specification versus the other.

However, the second approach suffers from a main weakness as indicated earlier; that is the technical inefficiency effects of different firms at any given time period, t , are equal to the same exponential function ($\exp[-\eta(t - T)] \equiv \exp[\eta(T - t)]$) of the corresponding firm-specific inefficiency effects at the last period of the panel (the u_{it} s). This implies that the ordering of the firms according to the magnitude of the technical inefficiency effects is the same at all time periods. Thus, the time-varying model of equation does not account for situations in which some firms may be relatively inefficient initially but become relatively more efficient in subsequent periods. (Furthermore, as Battese and Coelli (1995) indicated, a small error was detected in the first partial derivative with respect to η in the 1992 model of the program. This error would have only affected results when η was assumed to be non-zero).

Therefore, if the above two stages lead more or less to the same model specifications, we will take the efficiency estimates of the first stage which utilises the 1995 approach. However, if the two stages lead to different preferred model specifications, we will report the results of two stages and then compare the efficiency estimates result from each stage.

In the case of the cost function, the first stage leads us to select the Fourier-truncated without control variables but with efficiency correlates. The second stage leads us to select the translog-truncated without control variables as well. As such, it is plausible to assume that the inclusion of efficiency correlates in the first stage is the reason for the selection of the Fourier over translog in the first stage. Furthermore, as the second stage is estimated utilising Battese and Coelli's (1992) approach which does not allow us to include directly the efficiency correlates in the model and since there is no major differences between the specifications of the two stages, we will choose the result of **stage 1** as the cost preferred model; the Fourier-truncated model excluding control variables (capital, risk and time trend) but including all the efficiency correlates (the parameter estimates of the preferred model are shown in Table 8).

Table 8

Maximum likelihood estimates of the preferred cost function model

	The variables (all are logged)	Coefficient	Standard error	t-ratio
α		115.71	0.97	118.76
γ_1	lny1	0.54	0.54	1.00
γ_2	lny2	0.78	0.90	0.87
γ_3	lny3	0.17	0.38	0.44
β_1	Lnw1/w3	-14.15	0.65	-21.92
β_2	lnw2/w3	28.76	0.45	63.58

Table 8 (continuous)

	The variables (all are logged)	Coefficient	Standard error	t-ratio
γ_{11}	$\ln y_1 \ln y_1$	0.08	0.08	1.05
γ_{12}	$\ln y_1 \ln y_2$	-0.15	0.08	-1.77
γ_{13}	$\ln y_1 \ln y_3$	-0.05	0.08	-0.65
η_{11}	$\ln y_1 \ln w_1 / w_3$	0.07	0.19	0.38
η_{12}	$\ln y_1 \ln w_2 / w_3$	0.18	0.27	0.65
γ_{22}	$\ln y_2 \ln y_2$	0.01	0.13	0.09
γ_{23}	$\ln y_2 \ln y_3$	0.07	0.07	0.97
η_{21}	$\ln y_2 \ln w_1 / w_3$	0.02	0.24	0.08
η_{22}	$\ln y_2 \ln w_2 / w_3$	0.03	0.05	0.57
γ_{33}	$\ln y_3 \ln y_3$	-0.02	0.03	-0.59
η_{31}	$\ln y_3 \ln w_1 / w_3$	-0.01	0.14	-0.09
η_{32}	$\ln y_3 \ln w_2 / w_3$	-0.08	0.30	-0.27
β_{11}	$\ln w_1 / 3 \ln w_1 / w_3$	3.16	0.40	7.97
β_{12}	$\ln w_1 / w_3 \ln w_2 / w_3$	-1.69	0.36	-4.65
β_{22}	$\ln w_2 / w_3 \ln w_2 / w_3$	-16.62	0.38	-43.26
ϕ_1	$\cos(y_1)$	-0.19	0.27	-0.70
ω_1	$\sin(y_1)$	0.03	0.38	0.08
ϕ_2	$\cos(y_2)$	0.02	0.28	0.08
ω_2	$\sin(y_2)$	0.03	0.22	0.13
ϕ_3	$\cos(y_3)$	0.03	0.30	0.10
ω_3	$\sin(y_3)$	0.00	0.17	0.00
ϕ_4	$\cos(w_1 / w_3)$	-4.00	0.56	-7.10
ω_4	$\sin(w_1 / w_3)$	3.87	0.51	7.56
ϕ_5	$\cos(w_2 / w_3)$	-15.04	0.78	-19.18
ω_5	$\sin(w_2 / w_3)$	-14.05	0.76	-18.46
ϕ_{11}	$\cos(y_1 + y_1)$	0.00	0.02	-0.13
ω_{11}	$\sin(y_1 + y_1)$	-0.03	0.04	-0.68
ϕ_{12}	$\cos(y_1 + y_2)$	0.04	0.08	0.55
ω_{12}	$\sin(y_1 + y_2)$	-0.05	0.09	-0.54
ϕ_{13}	$\cos(y_1 + y_3)$	0.00	0.06	0.02
ω_{13}	$\sin(y_1 + y_3)$	0.00	0.04	0.11
ϕ_{14}	$\cos(y_1 + w_1 / w_3)$	-0.03	0.26	-0.12
ω_{14}	$\sin(y_1 + w_1 / w_3)$	0.08	0.12	0.63
ϕ_{15}	$\cos(y_1 + w_2 / w_3)$	0.05	0.21	0.24
ω_{15}	$\sin(y_1 + w_2 / w_3)$	-0.03	0.27	-0.10
ϕ_{22}	$\cos(y_2 + y_2)$	-0.01	0.07	-0.13
ω_{22}	$\sin(y_2 + y_2)$	0.04	0.01	5.96
ϕ_{23}	$\cos(y_2 + y_3)$	0.00	0.03	0.03
ω_{23}	$\sin(y_2 + y_3)$	0.00	0.04	-0.07
ϕ_{24}	$\cos(y_2 + w_1 / w_3)$	-0.01	0.20	-0.03
ω_{24}	$\sin(y_2 + w_1 / w_3)$	-0.10	0.16	-0.61
ϕ_{25}	$\cos(y_2 + w_2 / w_3)$	0.03	0.09	0.36
ω_{25}	$\sin(y_2 + w_2 / w_3)$	0.03	0.34	0.10
ϕ_{33}	$\cos(y_3 + y_3)$	0.01	0.00	1.67
ω_{33}	$\sin(y_3 + y_3)$	0.00	0.04	-0.02
ϕ_{34}	$\cos(y_3 + w_1 / w_3)$	-0.01	0.10	-0.13

Table 8 (continuous)

	The variables (all are logged)	Coefficient	Standard error	t-ratio
ω_{34}	$\text{Sin}(y_3+w_1/w_3)$	0.01	0.33	0.03
ϕ_{35}	$\text{Cos}(y_3+w_2/w_3)$	-0.02	0.20	-0.08
ω_{35}	$\text{Sin}(y_3+w_2/w_3)$	-0.02	0.14	-0.17
ϕ_{44}	$\text{Cos}(w_1/w_3+w_1/w_3)$	0.09	0.33	0.29
ω_{44}	$\text{Sin}(w_1/w_3+w_1/w_3)$	1.14	0.42	2.70
ϕ_{45}	$\text{Cos}(w_1/w_3+w_2/w_3)$	0.96	0.51	1.89
ω_{45}	$\text{Sin}(w_1/w_3+w_2/w_3)$	0.14	0.24	0.57
ϕ_{55}	$\text{Cos}(w_2/w_3+w_2/w_3)$	0.24	0.50	0.49
ω_{55}	$\text{Sin}(w_2/w_3+w_2/w_3)$	3.81	0.42	9.05
ϕ_{111}	$\text{Cos}(y_1+y_1+y_1)$	-0.01	0.05	-0.25
ω_{111}	$\text{Sin}(y_1+y_1+y_1)$	0.02	0.02	0.65
ϕ_{222}	$\text{Cos}(y_2+y_2+y_2)$	0.00	0.02	-0.21
ω_{222}	$\text{Sin}(y_2+y_2+y_2)$	0.00	0.03	-0.10
ϕ_{333}	$\text{Cos}(y_3+y_3+y_3)$	0.01	0.02	0.36
ω_{333}	$\text{Sin}(y_3+y_3+y_3)$	0.00	0.03	-0.06
ϕ_{444}	$\text{Cos}(w_1/w_3+w_1/w_3+w_1/w_3)$	0.33	0.17	1.90
ω_{444}	$\text{Sin}(w_1/w_3+w_1/w_3+w_1/w_3)$	0.23	0.22	1.01
ϕ_{555}	$\text{Cos}(w_2/w_3+w_2/w_3+w_2/w_3)$	0.32	0.28	1.11
ω_{555}	$\text{Sin}(w_2/w_3+w_2/w_3+w_2/w_3)$	-0.58	0.19	-2.99
δ_0		-0.05	0.57	-0.08
δ_1	L	0.13	0.56	0.23
δ_2	TA	0.00	0.00	0.34
δ_3	B	-0.09	0.23	-0.40
δ_4	J	0.13	0.69	0.18
δ_5	E	0.11	0.25	0.43
δ_6	Com	0.01	0.61	0.01
δ_7	Inv.	0.05	0.47	0.10
δ_8	Isl.	-0.06	0.39	-0.16
δ_9	3-FCR	-0.02	0.16	-0.12
δ_{10}	MS	-0.17	1.26	-0.14
sigma-squared (S)		0.08	0.01	9.42
gamma		0.008	0.006	1.263
Sigma-squared		0.001		
Sigma-squared (v)		0.082		
Lambda		0.089		
The relative contribution of the inefficiency effect to the total		0.003		
Log likelihood function		69.06		
LR test of the one-sided error		90.72		
[note that this statistic has a mixed chi-squared distribution]				

Source: Author's own estimation.

Estimated levels of cost efficiency

Efficiency estimates for the cost efficiency, derived from the preferred model, are summarised in Table 11 below.

Given the preferred cost function, efficiency estimates for banks in the countries under study averaged 95% and these estimates have slightly varied over time from 95% in 1992 to 94% in 2000. This suggests that the same level of output could be produced with approximately 95% of current inputs if banks under study were operating on the most efficient frontier. This level of inefficiency is somewhat less than the range of 10-15% for the 130 studies surveyed by Berger and Humphrey (1997)¹ and Berger and DeYoung (1997). These results are also less than the level of inefficiency found in European studies including Carbo et al.'s (2000) whose findings for a sample of banks, from twelve countries, show mean cost inefficiency of around 22 % for the period from 1989 to 1996.

Referring to Table 9, the average efficiency based on bank specialisation ranged from 93% for investment banks to 98% for Islamic banks. The efficiency scores based on geographical location, ranged from 89% in Jordan to 99% in Bahrain. Finally, based on asset size, the differences among technical efficiency scores are not significant where optimal bank size is between US\$ 2.5-5.0 billion and the largest banks seem to be somehow more efficient. These results are noticeably different from Carbo et al.'s (2000) findings on European savings banks who find that the least X-efficient banks were the largest in asset size.

Table 9

Cost efficiency in Jordan, Egypt, Saudi Arabia and Bahrain banking over 1992-2000

	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
Bahrain	100	100	100	100	100	99	99	99	99	99
Egypt	94	94	94	94	94	93	93	93	93	94
Jordan	90	89	89	89	89	89	89	88	88	89
Saudi Arabia	97	97	97	97	97	97	97	97	96	97
Commercial	95	95	95	95	94	94	94	94	94	94
Investment	93	93	93	93	93	93	93	93	93	93
Islamic	98	98	98	98	99	99	98	98	98	98
Other	97	96	96	96	96	96	96	96	96	96
All	95	95	95	95	95	94	94	94	94	95
Asset Size (US\$ million)										
	1-199	200-299	300-499	500-999	1,000-2,499	2,500-4,999	5,000-9,900	10000+		All
Bahrain	100	99	100	99	99	99	99	99		99
Egypt	95	94	94	94	94	93	92	90		94
Jordan	88	87	88	91	90			91		89
Saudi Arabia				98	98	98	98	95		97
All	95	93	94	95	95	96	96	94		95
Asset Size (US\$ million)										
	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
1-199.9	94	94	95	95	96	96	95	96	95	95
200-299	93	94	92	93	92	92	95	95	95	93
300-499	95	95	95	95	94	94	92	92	91	94
500-999	96	94	94	94	94	95	96	95	96	95
1,000-2,499	96	96	95	96	96	94	94	94	94	95
2,500-4,999	95	96	99	96	96	96	96	96	96	96
5,000-9,999	98	98	97	96	96	96	95	96	95	96
10000+	95	95	94	94	94	93	94	93	94	94
All	95	95	95	95	95	94	94	94	94	95

Source: Author's own estimation.

¹ Of these, 60 parametric studies found that the mean technical inefficiency is smaller than 15%.

To summarise the main findings, cost efficiency levels averaged around 95 percent over the period of 1992-2000 without noticeable change over the 1992-1999 period but have experienced a fall in 2000. Islamic banks are found to be the most cost efficient while investment banks are the least efficient. This result may partially explain the motives behind the increase in Islamic banking activities over the past few years; as the cost of funds for Islamic banks is relatively cheaper than the cost of funds for other financial institutions. On the other hand, intense competition between investment and commercial banks might explain the competitive disadvantages of the investment banks in terms of their market share and expose the motives for increased mergers and consolidation activity between such banks.

Based on assets size, large banks seem to be relatively more cost efficient, in general. This result suggests that large banks enjoy several advantages compared to small banks. These include the ability of large banks to utilise more efficient technology with less cost, the ability of these banks to set up more specialised staff for the most profitable activities and the ability of these banks to provide better quality output and therefore charge higher prices. Geographically, Bahrain is the most cost and profit efficient banking systems while Jordan is the least cost and profit efficient.

Finally, while the countries under study have implemented many economic and financial reforms over the last twenty years or so as indicated earlier, these reforms do not appear to have had much impact on banking sector efficiency. Given our findings, it seems that more reform may be needed to improve their efficiency. Perhaps the move to create a single GCC market may help to facilitate these developments as the creation of a similar European Single market appears to have had a positive impact on European bank efficiency (see European Commission (1997)).

Estimated levels of scale elasticities

Productive efficiency requires optimising behaviour with respect to outputs as well as inputs as indicated earlier. Regarding outputs, optimal behaviour relates to producing the level of outputs that correspond to the lowest cost per unit. For the cost function, the optimal output level is possible if economies and diseconomies exist at different output levels; that is at some point, there will be constant returns defining the optimal level of production. Economies of scale exist if, over a given range of output, per unit costs decline as output increases. Increases in per unit cost correspond to decreasing returns to scale. A scale efficient firm will produce where there are constant returns to scale; that is, changes in output result in proportional changes in costs (Evanoff and Israilevich, 1991).

Given the cost function specification, the scale economy measure is a cost elasticity; the percent change in cost with respect to a percent change in output. On this basis, the results suggest existence of scale diseconomies across the banks under study and the scale diseconomies for these banks ranged from around 3% in 1992 to 6% in 2000 and averaged 5% over the 1992-2000 period (Table 10 shows the details¹). Thus, a 100 percent increase in the level of outputs would lead to about 105% percent increase in total costs. The magnitude of these scale diseconomies estimates is not different from other banking literature that finds evidence of diseconomies in the US banking market. For example, see Berger et al. (1993), Hughes et al. (1995) and McAllister and McManus (1993).

Based on the size of banks' assets, the optimal bank size are those in the ranges of US\$ 5-10 billion where banks in this category experience increasing returns to scale. In addition, scale economies increase with size, and optimal bank size is inexhaustible which supports an argument for further consolidation. Based on geographical location, Saudi Arabian and (to a lesser extent) Egyptian banks seem to have the largest unrealised scale economies (see Table 6.12 for details).

¹ See Table 10's footnote to observe if these values are statistically significant from unity.

Table 10

Scale elasticities in the banking sectors of Jordan, Egypt, Saudi Arabia and Bahrain over 1992-2000

	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
Bahrain	1.23	1.25	1.22	1.24	1.27	1.22	1.22	1.21	1.26	1.23
Egypt	0.92	0.97	0.92	0.90	0.96	1.00	1.02	1.03	1.00	0.97
Jordan	1.14	1.09	1.16	1.23	1.21	1.20	1.17	1.15	1.07	1.16
Saudi Arabia	0.94	0.90	0.88	0.92	0.89	0.92	0.90	0.93	0.97	0.92
Commercial	0.94	0.95	0.92	0.92	0.94	0.95	0.95	0.95	0.94	0.94
Investment	1.15	1.15	1.16	1.24	1.22	1.27	1.31	1.28	1.30	1.23
Islamic	1.19	1.30	1.34	1.40	1.49	1.42	1.39	1.31	1.29	1.35
Other	1.26	1.29	1.25	1.17	1.25	1.27	1.24	1.32	1.29	1.26
All	1.03	1.04	1.02	1.03	1.06	1.07	1.07	1.07	1.06	1.05
Asset Size (US\$ million)										
	1-199	200-299	300-499	500-999	1,000-2,499	2,500-4,999	5,000-9,900	10000+		All
Bahrain	1.33	1.15	1.25	1.38	1.42	1.23	1.15	0.46		1.23
Egypt	0.79	0.88	0.92	0.97	1.17	1.15	0.97	0.67		0.97
Jordan	1.06	1.15	1.15	1.25	1.29			0.90		1.16
Saudi Arabia				0.83	1.03	1.15	0.95	0.69		0.92
All	1.05	1.01	1.06	1.13	1.19	1.16	0.98	0.67		1.05
Asset Size (US\$ million)										
	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
1-199.9	1.01	1.11	1.10	1.03	1.09	0.98	0.94	1.06	1.03	1.05
200-299	1.01	1.08	0.92	1.05	1.02	1.07	1.07	0.93	0.81	1.01
300-499	1.07	1.04	1.06	1.09	1.08	1.10	1.01	1.02	1.02	1.06
500-999	1.09	1.02	1.00	1.09	1.16	1.07	1.24	1.19	1.18	1.13
1,000-2,499	1.05	1.10	1.19	1.14	1.19	1.29	1.26	1.23	1.23	1.19
2,500-4,999	1.13	1.05	0.94	1.10	1.05	1.06	1.20	1.33	1.33	1.16
5,000-9,999	0.99	0.96	0.99	0.84	0.97	0.99	1.06	1.01	1.04	0.98
10000+	0.90	0.81	0.73	0.69	0.66	0.57	0.62	0.55	0.62	0.67
All	1.03	1.04	1.02	1.03	1.06	1.07	1.07	1.07	1.06	1.05

Note: The scores that fall within the ranges [0.983-1.016] and [0.966-1.033] are not statistically different from one at 5 percent and 1 percent level respectively for two-tailed test.

Source: Author's own estimation.

To summarise, (cost) scale elasticity estimate for the banking systems under study is around 105% and this did not noticeably change over 1992-2000. This implies that increasing the size of operations by 100 percent results in an increase in cost by 105 percent. In other words, scale diseconomies predominate. Nevertheless, we do not find evidence of significant scale economies for the largest banks in the sample. Overall, it appears that scale elasticities are most prevalent for commercial banks and for the largest banks in general.

Estimated levels of scale efficiency

The scale elasticity measure, as indicated earlier, is an elasticity associated with a particular output level and indicates the relative change in cost associated with an increment change from this output level. Scale inefficiency (I), on other hand, can be measured as the aggregate cost of F inefficient firms ($\varepsilon \neq 1.0$) relative to the cost of a single efficient firm ($\varepsilon = 1.0$).

Given the following representation for the cost function: $\ln C = a + b (\ln Y) + .5 c (\ln Y)^2$, then the scale elasticity for inefficient firms $= \varepsilon_I = \partial \ln C_I / \partial \ln Y_I = b$. On this basis, scale inefficiency can be written as: $I = e^{(.5/c)(1-\varepsilon_I)^2} - 1.0$, that is scale inefficiency is a function of the

first and second derivatives of the function with respect to output (the second derivation helps to reach c which is the key for calculation of inefficiency). Note, if the estimated scale elasticity is insignificantly different from unity, this does not imply scale inefficiency is insignificantly different from zero because the statistical difference of the elasticity measure from a value of unity depends entirely on the standard error of the estimated coefficient b.

Given the cost function specification of the stochastic frontier, scale efficiency averaged around 65% for banks under study over 1992 to 2000. Furthermore, there is a significant drop in scale efficiency over time when it decreased from around 72% in 1992 to reach 60% percent in 2000. According to geographical location, the efficiency scores ranged from 72% for Jordan and Saudi Arabian banks to 51% for Bahrain banks. Furthermore, commercial banks are the most efficient with cost efficiencies around 70% while the least efficient are the Islamic banks (Table 11). Furthermore, the results generally show that some categories of small and large banks are scale efficient while other ranges do have similar efficiency levels.

Table 11

Cost Scale inefficiency for the banking sectors of Jordan, Egypt, Saudi Arabia and Bahrain over 1992-2000

	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
Bahrain	47	49	44	51	53	53	49	46	52	49
Egypt	24	24	31	33	32	35	36	41	40	33
Jordan	21	26	25	34	31	34	30	27	25	28
Saudi Arabia	20	21	27	27	23	29	30	36	40	28
Commercial	24	25	30	34	29	31	30	33	33	30
Investment	32	30	30	42	39	44	45	40	42	38
Islamic	34	47	50	59	74	77	71	65	55	59
Other	38	43	31	25	38	46	46	57	70	44
All	28	29	32	36	35	38	37	39	40	35
Asset Size (US\$ million)										
	1-199	200-299	300-499	500-999	1,000-2,499	2,500-4,999	5,000-9,900	10000+		All
Bahrain	44	27	41	54	75	54	17	79		49
Egypt	44	26	17	28	51	23	28	50		33
Jordan	21	20	21	39	47			20		28
Saudi Arabia				25	26	27	16	43		28
All	38	24	24	37	49	31	19	48		35
Asset Size (US\$ million)										

Table 11 (continuous)

	1992	1993	1994	1995	1996	1997	1998	1999	2000	All
1-199.9	30	32	39	37	40	46	46	46	51	38
200-299	20	33	19	28	25	24	24	16	26	24
300-499	25	19	19	29	33	35	21	20	14	24
500-999	30	35	33	40	37	35	44	35	37	37
1,000-2,499	42	47	49	47	53	55	50	51	47	49
2,500-4,999	25	16	54	24	3	13	28	48	49	31
5,000-9,999	10	10	19	37	20	22	12	14	23	19
10000+	30	29	40	43	46	52	50	69	56	48
All	28	29	32	36	35	38	37	39	40	35

Source: Author's own estimation.

6. Conclusion

A major aim of this study is to estimate efficiency levels in various Arabic banking sectors by applying various statistical analyses to a data set on Jordan, Egypt, Saudi Arabia and Bahrain. This study employs cost efficiency concept using a number of different measurement methods (including the stochastic frontier approach, specification of the Fourier-flexible functional form versus the translog form, and inclusion of bank's asset quality and financial capital in a number of different ways) to a single data set.

In choosing the 'preferred' cost model, we follow the recent efficiency methodologies that proceed by testing various model specifications to arrive at the preferred model. Based on the preferred models, cost efficiency measures are reported for the banks in the countries under study. Given cost efficiency, the preferred model is the Fourier-truncated form that excludes the control variables (capital adequacy, asset quality and the time trend) but includes all the environmental variables.

Based on the chosen preferred model, cost efficiency averaged around 95% over the 1992-2000 period. Islamic banks are found to be the most cost efficient, while investment banks are the least. Based on bank asset size, large banks seem to be relatively more cost efficient. Geographically, Bahrain is the most cost efficient while Jordan is the least. It should be noted that these results, in general, are similar to those found in other US and European banking studies.

Based on the estimated preferred model, we also report scale elasticity and scale efficiency measures for the banks under study. The cost scale elasticity estimates reveal diseconomies of around five percent and the cost scale inefficiency estimates also suggest that banks are 65% scale efficient. Islamic and commercial banks are again found to be the most cost scale efficient. Large banks are also generally found to be more efficient than smaller institutions. In addition, geographically, Saudi Arabian and Egyptian banks seem to be the most cost scale efficient.

A major finding of this study is that there is little evidence to suggest that the major economic and financial reforms undertaken in Jordan, Egypt, Saudi Arabia and Bahrain over the last decade have had a noticeable impact on improvement in banking sector efficiency. The main policy recommendation from this study, therefore, is that these countries need to continue the reform process in order to enhance financial sector performance.

References

1. Aigner, D., C. Lovell, and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, 6 (May), pp. 21-37.
2. Allen, R. and A. Rai (1996), "Operational Efficiency in Banking: An International Comparison", *Journal of Banking & Finance*, pp. 655-672.
3. Altunbas, Y., E. Gardener, and P. Molyneux (1996), "Cost Economies and Efficiency in EU Banking Systems", *Journal of Economics and Business*, 48, pp. 217-230.
4. Altunbas, Y., L. Evans, and P. Molyneux (1997), "Bank Ownership and Efficiency", *Institute of European Finance, University of Wales, Bangor, Unpublished*, (WP 18/97).
5. Altunbas, Y., E. Gardener, P. Molyneux, and B. Moore (1998), "Efficiency in European Banking", *University of Wales, Bangor, Un-published*, (WP 2/98).
6. Altunbas, Y., M. Liu, P. Molyneux and R. Seth (1999), "Efficiency and risk in Japanese banking", *Journal of Banking & Finance*, 24, pp. 1605-1628.
7. Aly, H., R. Grabowski, C. Pasurka, and N. Rangan (1990), "Technical, Scale, and Allocative Efficiencies in U.S. Banking: An Empirical Investigation", *The Review of Economics and Statistics*, 72, pp. 211-218.
8. Barr, R., Seiford, L., and Siems, T. (1994), "Forecasting bank failure: A non-parametric approach", *Recherches Economiques de Louvain*, 60, 411-29.
9. Battese, G. and T. Coelli (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", *The Journal of Productivity Analysis*, 3, pp. 153-169.
10. Battese, G. and T. Coelli (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", *Empirical Economics*, 20, pp. 325-332.
11. Bauer, P. and D. Hancock (1993), "The efficiency of the Federal Reserve in providing check processing services", *Journal of Banking and Finance*, 17, pp. 287-311.
12. Benston, G., G. Hanweck, and D. Humphrey (1982a), "Scale economies in banking: a restructuring and reassessment", *Journal of Money, Credit and Banking*, 14, pp. 435-456.
13. Benston, G., G. Hanweck, and D. Humphrey, (1982b), "Operating costs in commercial banking", *Federal Reserve Bank of Atlanta*, *Economic Review*, pp. 6-21.
14. Berg, S., F. Forsund and N. Bukh (1995), "Banking efficiency in the Nordic countries: a few-country Malmquist index analysis" (mimeo).
15. Berg, S., F. Forsund, and E. Jansen, (1992), "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89", *Scand. J. of Economics*, 94, Supplement, pp. 211-228.
16. Bergendahl G. (1998), "DEA and Benchmarks- An Application to Nordic Banks", *Annals of Operations Research*, Vol. 82, pp. 233-249.
17. Berger, A. (1995), "The profit-structure relationship in banking-tests of market power and efficient-structure hypothesis", *Journal of Money, Credit and Banking*, 27(2), pp. 404-431.
18. Berger, A. and D. Humphrey (1992), "Measurement and Efficiency Issues in Commercial Banking", in Zvi Griliches (Eds.), *Output Measurement in the Services Sectors*, National Bureau of Economic Research Studies in Income and Wealth, University of Chicago, Chicago, 56, pp. 245-279.
19. Berger, A. and L. Mester (1997), "Inside the black box: What explains differences in the efficiencies of financial institutions?", *Journal of Banking & Finance*, 21, pp. 895-947.
20. Berger, A. and L. Mester (1999), "What Explains the Dramatic Changes in Cost and Profit Performance of the U.S. Banking Industry", *Federal Reserve Board, Washington, D.C.*, (13/1999).
21. Berger, A. and R. DeYoung (1997), "Problem Loans and Cost Efficiency in Commercial Banks", *Federal Reserve Board, Washington, D.C.*, (WP 8/1997).
22. Berger, A. and T. Hannan (1995), "The Efficiency Cost of Market Power in the Banking Industry: A Test of the Quiet Life and Related Hypotheses", *The Review of Economics and Statistics*, pp. 454-465.

23. Berger, A., G. Hanweck, and D. Humphrey (1993a), "Bank efficiency derived from the profit function", *Journal of Banking & Finance*, 17, pp. 317-347.
24. Berger, A., Leusner, J. and J. Mingo (1997b), "The efficiency of bank branches", *Journal of Monetary Economics*, 40.
25. Berger, A., W. Hunter, and S. Timme (1993b), "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future", *Journal of Banking & Finance*, 17, pp. 221-249.
26. Bhattacharya, A., C. Lovell, and P. Sahay (1997), "The Impact of Liberalisation on the Productive Efficiency of Indian Banks", *European Journal of Operational Research*, (98), pp. 332-345.
27. Bhattacharya, S. and A. Thakor (1992), "Contemporary Banking Theory", *Journal of Financial Intermediation* 3 (October), pp. 2-50.
28. Carbo, S., E. Gardener, and J. William (1999), "A Note on Technical Change in Banking: The Case of European Savings Banks", *Institute of European Finance, University of Wales, Bangor, Un-published*.
29. Casu, B. and P. Molyneux (2001), "Efficiency in European Banking", 'in Goddard, J., P. Molyneux, J. Wilson (2001), *European Banking: Efficiency, Technology and Growth*, John Wiley: London, pp. 121-179.
30. Cebenoyan, A., Cooperman, E., Register, C. (1993), "Firm inefficiency and the regulatory closure of S& Ls: An empirical investigation", *Review of Economics and Statistics*, 75, pp. 540-45.
31. Chaffai, M., Dietsch, M. (1995), "Should banks be 'universal'? The relationship between economies of scope and efficiency in the French banking industry", *Working paper, University of Robert Schuman of Strasbourg, France*.
32. Chalfant, J. and A. Gallant (1985), "Estimating substitution elasticities with the Fourier cost function", *Journal of Econometrics*, 28, pp. 205-222.
33. Chang, C., Hasan, I., Hunter, W. (1998), "Efficiency of multinational banks: An empirical investigation", *Applied Financial Economics*, 8, pp. 1-8.
34. Clark, J. (1988), "Economies of Scale and Scope At Depository Financial Institutions: A Review of the Literature", *Federal Reserve Bank of Kansas City, Economic Review*, 73, pp. 16-33.
35. Clark, J. (1996), "Economic Cost, Scale Efficiency, and Competitive Viability in Banking", *Journal of Money, Credit and Banking*, 28 (3).
36. Coelli, T. (1996b), "A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation", *Centre for Efficiency and Productivity Analysis (CEPA), Department of Econometrics, University of New England, Armidale, Australia*, (WP 7/96).
37. Coelli, T. (1996c), "Measurement and Source of Technical Efficiency in Australian Coal-fired Electricity Generation", *Centre for Efficiency and Productivity Analysis (CEPA), Department of Econometrics, University of New England, Armidale, Australia*, (WP 1/96).
38. Dietsch, M., and Lozano-Vivas, A. (2000), "How the environment determines banking efficiency: A comparison between French and Spanish industries", *Journal of Banking and Finance*, 24, pp. 985-1004.
39. Eastwood, B. and A. Gallant (1991), "Adaptive rules for semi-nonparametric estimators that achieve asymptotic normality" *Economic Theory*, 7, pp. 307-340.
40. Elbadawi, I., A. Gallant and G. Souza (1983), "An elasticity can be estimated consistently without a priori knowledge of functional form", *Econometrica*, 51, pp. 1731-1753.
41. Elyasiani, E. and S. Mehdiian (1990a), "Efficiency in commercial banking industry, a production frontier approach", *Applied Economics*, 22, pp. 539-551.
42. Elyasiani, E. and S. Mehdiian (1990b), "A non-parametric approach to measurement of efficiency and technological change: the case of large US banks", *Journal of Financial Services Research*, 4, pp. 157-168.
43. Elyasiani, E. and S. Mehdiian (1995), "The comparative efficiency performance of small and large US commercial banks in the pre-and post-deregulation eras", *Applied Economics*, 27, pp. 1069-1079.

44. European Commission (1997), "The Single Market Review" 'Credit Institutions and Banking, Subseries II: Impact on Services', Vol. 3, (London: Kogan Page).
45. Evanoff, D. and R. Israilevich (1991), "Productive efficiency in banking", 'Economic Perspective, Federal Reserve of Chicago', 15 (July/August), pp. 11-32.
46. Evanoff, D. and R. Israilevich (1995), "Scale Elasticity Versus Scale Efficiency in Banking", 'Southern Journal of Economics', 61, pp. 1036-1047. Evanoff, D. and R. Israilevich (1995), "Scale Elasticity Versus Scale Efficiency in Banking", 'Southern Journal of Economics', 61, pp. 1036-1047.
47. Ferrier, G., Grosskopf, S., Hayes, K., and Yaisawarng, S. (1993), "Economies of diversification in banking industry: A frontier approach", 'Journal of Monetary Economics', 31, pp. 229-49.
48. Gallant, A. (1981), "On the bias in Flexible Functional forms and essentially unbiased form: The Fourier Flexible form", 'Journal of Econometrics', 15, 211-245.
49. Gallant, A. (1982), "Unbiased determination of production technologies", 'Journal of Econometrics', 20, pp. 285-324.
50. Gallant, A. and G. Souza (1991), "On the asymptotic normality of Fourier Flexible form estimates", 'Journal of Econometrics' 50, pp. 329-353.
51. Greene, W. (1990), "A Gamma-distributed Stochastic Frontier Model", 'Journal of Econometrics', 46, pp. 141-164.
52. Greenbaum, S. (1967), "A study of bank cost", 'National Banking Review', June, pp. 415-434.
53. Griffell-Tatjé, E., and C. Lovell (1997), "Profits and Productivity", 'The Wharton School, University of Pennsylvania', (WP 18/97).
54. Hermalin, B., Wallace, N. (1994), "The determinants of efficiency and solvency in savings and loans", 'Rand Journal of Economics', 25, pp. 361-381.
55. Hughes, J., Lang, W., Mester, L. and C. Moon (1995), "Recovering banking technologies when managers are not risk-neutral", 'Conference on Bank Structure and Competition, Federal Reserve Board of Chicago', May, pp. 349-368.
56. Hughes, J., Mester, L. and C. Moon (1996a), "Efficient banking under interstate branching", 'Journal of Money, Credit and Banking', 28, pp. 1045-1071.
57. Hughes, J., Mester, L. and C. Moon (1996b), "Safety in numbers? Geographic diversification and bank insolvency risk", 'Working paper 96-14, Federal Reserve Bank of Philadelphia'.
58. Hughes, J., Mester, L. and C. Moon (1997), "Recovering risky technologies using the almost ideal demand system: An application to U.S. banking", 'Working paper 97-8, Federal Reserve Bank of Philadelphia'.
59. Humphrey, D. and L. Pulley (1997), "Banks' Responses to Deregulation: Profits, Technology, and Efficiency", 'Journal of Money, Credit and Banking', 29(1), pp. 73-93.
60. Intarachote, T. (2000), "Financial liberalisation in Thailand", 'PhD, University of Wales, Bangor'.
61. Ivaldi, M., Ladoux, N., Ossard, H. and M. Simioni (1996), "Comparing Fourier and Translog specification of multiproduct technology: evidence from an incomplete panel of French farmers", 'Journal of Applied Economics', Vol. 11, pp. 649-667.
62. Jagtiani, J. and A. Khanthavit (1996), "Scale and scope economies at large banks: Including off-balance sheet products and regulatory effects (1984-1991)", 'Journal of Banking & Finance', 20, pp. 1271-1287.
63. Jondrow, J., C. Lovell, I. Materov, and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in The Stochastic Frontier Production Function Model", 'Journal of Econometrics', 19, pp. 233-238.
64. Kaparakis, E., Miller, S., and Noulas, A. (1994), "Short-run cost inefficiency of commercial banks: A flexible stochastic frontier approach", 'Journal of Money, Credit, and Banking', 26, pp. 875-93.
65. Kumbhakar, S., S. Ghosh and J. McGukin (1991), "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms", 'Journal of Business and Economic Statistics', 9, pp. 279-286.

66. Kwan, S., Eisenbeis, R. (1994), "An analysis of inefficiencies in banking: A stochastic cost frontier approach", 'Federal Reserve Bank of San Francisco, USA', Working paper (December).
67. Leightner, E. and C. Lovell (1998), "The Impact of Financial Liberalisation on the Performance of Thai Banks", 'Journal of Economics and Business', Vol. 50, pp. 115-131.
68. McAllister, R. and D. McManus (1993), "Resolving the scale efficiency puzzle in banking", 'Journal of Banking & Finance', 17, pp. 389-405.
69. McKinnon, R. (1973), "Money and Capital in Economic Development". 'Washington, D.C.: Brookings Institution'.
70. Meeusen, W. and J. van den Broeck (1977), "Efficiency estimation from Cobb-Douglas production functions with composed error", 'International Economic Review', 18, pp. 435-444.
71. Mester, L. (1993), "Efficiency in the savings and loan industry", 'Journal of Banking & Finance', 17, pp. 267-286.
72. Mester, L. (1996), "A Study of Bank Efficiency Taking into Account Risk-Preferences", 'Journal of Banking & Finance', 20, pp. 389-405.
73. Mester, L. (1996), "Measuring Efficiency at U. S. Banks: Accounting for Heterogeneity is Important", 'European Journal of Operational Research', 98, pp. 230-243.
74. Mitchell, K. and N. Onvural (1996), "Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier-Flexible Functional Form", 'Journal of Money, Credit and Banking', 28(2), pp. 178-199.
75. Pastor, J. (1995), "How to Account for Environmental Effects in DEA: an Application to Bank Branches", 'Working paper, University of Alicante, Spain'.
76. Pi, L., Timme, S., (1993), "Corporate control and bank efficiency", 'Journal of Banking and Finance', 17, pp. 515-530.
77. Pitt, M. and L. Lee (1981), "Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry", 'Journal of Development Economics', 9, 43-64.
78. Sealey, C. and J. Lindley (1977), "Inputs, outputs and a theory of production and cost at depository financial institutions", 'Journal of Finance', 32, pp. 1251-1266. Sealey, C. and J. Lindley (1977), "Inputs, outputs and a theory of production and cost at depository financial institutions", 'Journal of Finance', 32, pp. 1251-1266.
79. Shaw, E. (1973), "Financial Deepening in Economic Development", 'New York: Oxford University Press'.
80. Sinkey, J. (1992), "Commercial Bank Financial Management In the Financial-Services Industry", 'Macmillan Publishing Company, New York', Fourth Edition.
81. Spong, K., R. Sullivan and R. DeYoung (1995), "What makes a bank efficient? A look at financial characteristics and bank management and ownership structure", 'Financial Industry Perspective, Federal Reserve Bank of Kansas City'.
82. Tolstov, G. (1962), "Fourier Series", 'London: Prentice-Hall Int.'.
83. Vennet, R. (1998), "Cost and profit dynamics in financial conglomerates and universal banks in Europe", 'paper presented at the SUERF/CFS Colloquim, Frankfurt, 15-17 October'.
84. Whalen, G. (1991), "A proportional hazards model of bank failure: An examination of its usefulness as an early warning tool", 'Federal Reserve Bank of Cleveland, Economic Review', Quarter 1.
85. Wheelock, D. and P. Wilson (1995), "Explaining bank failures: Deposit insurance, regulation, and efficiency", 'Review of Economics and Statistics', 77, pp. 689-700.
86. White, H. (1980), "A heteroscedasticity consistent covariance matrix estimator and a direct test of heteroscedasticity", 'Econometrica', 48, pp. 817-818.
87. William, J. and E. Gardener (2000), "Efficiency and European Regional Banking", 'University of Wales, Bangor, Un-published'.
88. Zaim, O. (1995), "The effect of financial liberalisation on the efficiency of Turkish commercial banks", 'Applied financial Economics', 5, pp. 257-264.