

# UMA AVALIAÇÃO DA EFICIÊNCIA DOS BANCOS PORTUGUESES\*

Miguel Boucinha\*\*

Nuno Ribeiro\*\*

Thomas Weyman-Jones\*\*\*

## 1. INTRODUÇÃO

O sistema bancário desempenha um papel fulcral no sistema financeiro bem como na economia real, tal como evidenciado no contexto da crise financeira de 2008. O seu funcionamento adequado reveste-se de especial importância na intermediação de fundos na economia, ao mesmo tempo que disponibiliza uma vasta gama de serviços financeiros. Para que tal se verifique, os bancos devem não só fazer uma gestão adequada do risco, mas também alocar os seus recursos de uma forma tão eficiente quanto possível. Assim, a avaliação do desempenho dos bancos constitui um tópico de grande importância e que tem merecido especial atenção na literatura relacionada.

O objetivo do presente artigo é analisar a evolução do desempenho do sistema bancário português entre 1992 e 2004, um período em que o mesmo foi alvo de substanciais alterações, incluindo o processo de liberalização, consolidação e inovação financeira<sup>1</sup>. Estes desenvolvimentos surtiram um impacto profundo na estrutura de mercado e na tecnologia dos bancos e, através da análise de uma função custo, pretende-se avaliar de que forma afetaram os custos marginais e a produtividade total dos bancos, que é decomposta no efeito de alterações na eficiência de escala e na eficiência-custo, assim como no efeito do progresso tecnológico<sup>2</sup>. Assim, este procedimento permite não só quantificar o crescimento da produtividade total de fatores, mas também identificar se as alterações na produtividade foram motivadas por um movimento para um ponto diferente da função custo, por uma aproximação à fronteira de custo, ou por deslocamentos da própria fronteira.

Entre os trabalhos empíricos anteriores sobre a eficiência dos bancos portugueses incluem-se Mendes e Rebelo (1999), Mendes e Rebelo (2000), Pinho (2001), Canhoto e Dermine (2003), Lima (2008) e Lima e Pinho (2008). A maioria destes estudos recorrem à estimação de funções custo *translog* através da técnica de análise de fronteiras estocásticas (*Stochastic Frontier Analysis – SFA*), ao passo que a abordagem seguida por Canhoto e Dermine (2003) assenta na estimação de uma fronteira não paramétrica estimada por *Data Envelopment Analysis – DEA*. Por sua vez, o estudo de Mendes e Rebelo (2000) reporta resultados obtidos através de ambas as metodologias.

\* As opiniões expressas no artigo são da responsabilidade dos autores, não coincidindo necessariamente com as do Banco de Portugal ou do Eurosistema. Eventuais erros e omissões são da exclusiva responsabilidade dos autores.

\*\* Banco de Portugal, Departamento de Estudos Económicos.

\*\*\* Department of Economics, Loughborough University, UK

(1) Para uma breve descrição do processo de liberalização ver Ribeiro (2007) e Banco de Portugal (2009).

(2) A eficiência-custo é frequentemente designada na literatura como eficiência-X.

Embora os estudos referidos recorram a diferentes abordagens teóricas e empíricas e cubram diferentes períodos amostrais (compreendidos entre 1987 e 2004), todos à exceção de Mendes e Rebelo (1999) encontram evidência de ganhos de produtividade por parte dos bancos portugueses. Contudo, e tal como seria de esperar dada a diversidade das abordagens, os referidos estudos obtêm estimativas diferentes para os níveis de ineficiência-X, divergindo também na identificação de alterações nas melhores práticas, a que correspondem deslocações da fronteira eficiente, e na distância a que os bancos operam face à mesma. De facto, os estudos que assumem que a fronteira é constante ao longo do tempo não permitem decompor estes dois últimos efeitos, sendo que, nesses casos, todas as alterações na produtividade são atribuídas a alterações na eficiência-custo. Alguns estudos documentam diferenças na eficiência tendo em conta a heterogeneidade entre bancos. Assim, Pinho (2001) e Mendes e Rebelo (2000) concluem que os bancos de capital público tendem a ter um desempenho inferior à média, ao passo que os resultados obtidos em Canhoto e Dermine (2003) indicam que os bancos criados após 1984 e os bancos estrangeiros têm um melhor desempenho do que os bancos mais antigos (sendo que estes incluem os bancos públicos), que desenvolveram a sua atividade sob as condições de mercado fortemente restritivas que prevaleceram anteriormente. Finalmente, Mendes e Rebelo (2000) e Lima (2008) encontram evidência de que as fusões entre bancos contribuíram para uma melhoria do seu desempenho.

A restante parte deste artigo encontra-se organizada da seguinte forma. Na Secção 2 é apresentada a metodologia e os dados utilizados e na Secção 3 são expostos os resultados empíricos do artigo. Esta secção encontra-se dividida em 6 subsecções nas quais se inclui a discussão das estimativas para os custos marginais dos bancos portugueses, para o custo sombra dos capitais próprios, para a eficiência de escala, para a eficiência-custo, para o progresso tecnológico e para a produtividade total dos fatores. Por fim, a secção 4 apresenta as principais conclusões da análise.

## 2. METODOLOGIA E DADOS

A forma de modelar a atividade dos bancos tem sido alvo de debate na literatura, essencialmente devido a posições divergentes no que respeita à classificação dos depósitos de clientes como *inputs* ou como *outputs*. Por um lado, a denominada *abordagem da produção* considera os bancos como empresas que produzem serviços relacionados com os depósitos e empréstimos, pelo que identificam como *output* o número de depósitos sob gestão e o número de empréstimos originados e como *inputs* o trabalho e o capital físico. Por outro lado, de acordo com a *abordagem da intermediação* (Sealey e Lindley (1977)), a principal atividade dos bancos consiste na concessão de empréstimos e no investimento em títulos e outros ativos com base em fundos obtidos através da captação de depósitos e dos mercados de dívida por grosso, bem como em capitais próprios<sup>3</sup>. Existem argumentos teóricos válidos a favor de ambas as abordagens não havendo, da mesma forma, uma clara preferência por nenhuma delas nas aplicações empíricas<sup>4</sup>. Contudo, tal como destacado por Hughes,

(3) Para mais detalhes acerca das denominadas abordagem da produção e abordagem da intermediação, ver Freixas e Rochet (1998), pp. 77-79.

(4) O facto da abordagem da produção identificar como *outputs* o número dos empréstimos originados e dos depósitos sob gestão constitui uma complicação adicional devido a problemas na disponibilidade dos dados necessários. Os estudos que seguem esta abordagem tendem a contornar o problema aproximando o número dos empréstimos e depósitos pelo seu valor.

Mester e Moon (2001), deve evitar-se considerar os depósitos tanto como *input* como *output*, uma vez que neste caso, por um lado, a função custo teria de incluir o nível dos depósitos (devido à sua natureza de *output*) e, por outro lado, teria de incluir também o preço dos mesmos ao mesmo tempo que a definição dos custos deveria incluir as despesas relacionadas com os depósitos (devido à sua natureza de *input*). Neste caso, a escolha ótima de um *input* – os depósitos – não dependeria do preço do mesmo, uma vez que a sua quantidade estaria a ser mantida constante.

Existem essencialmente dois motivos pelos quais a captação de depósitos é uma atividade atrativa para os bancos. Por um lado, tal como sugerido pela abordagem da produção, os depósitos geram proveitos relacionados com comissões e constituem um produto que cria valor por si mesmo, uma vez que o público em geral não tem acesso às mesmas oportunidades de investimento dos bancos. Por outro lado, os depósitos constituem uma fonte de financiamento relativamente barata e estável. De qualquer das formas, os bancos dedicam uma parte considerável dos seus recursos à captação e gestão dos depósitos. No entanto, embora em menor grau, a aquisição e gestão de qualquer *input* tem custos associados.

Com base nos resultados do teste proposto em Hughes e Mester (1993) optou-se neste artigo por considerar os depósitos como um *input*, uma vez que a elasticidade dos custos totais deduzida dos juros pagos sobre depósitos em relação ao nível dos depósitos é negativa. A ideia subjacente a este teste é de que um aumento na utilização de um *input* deverá levar a uma diminuição dos custos relacionados com os outros *inputs*, ao passo que um aumento na produção de um *output* deverá surgir associado a um aumento de custos. A especificação adotada tem a vantagem adicional de permitir uma definição mais completa dos custos dos bancos, uma vez que a abordagem alternativa ignora por completo os custos de financiamento, limitando assim a análise aos custos operacionais. Esta limitação poderia enviesar os resultados do estudo uma vez que alguns bancos poderão estar dispostos a suportar custos operacionais mais elevados (com mão-de-obra e equipamento) com o objetivo de otimizar a sua estrutura de financiamento de forma a minimizar os custos financeiros.

A inclusão dos capitais próprios como um *input* na função custo justifica-se por um argumento semelhante pois, tal como destacado por Hughes, Mester e Moon (2001), caso esta variável fosse omitida poderiam ser obtidos resultados enviesados uma vez que os bancos cuja estrutura de financiamento tem um maior peso de capitais próprios em detrimento de dívida iriam espuriamente surgir como mais eficientes. O facto de os capitais próprios serem tratados como um *input* fixo e não variável é justificado por limitações à escolha do nível ótimo de capital relacionadas com restrições regulamentares e com objetivos de *rating*/reputação. Adicionalmente, devido aos custos associados às operações de emissão de capital, os bancos tendem a emitir capital em tranches relativamente elevadas, pelo que o nível de capital num dado momento deverá ser suficiente não apenas para cobrir os riscos atuais, mas também para acomodar aumentos de risco, possivelmente decorrentes do crescimento dos ativos, que venham a ocorrer no futuro. Desta forma, os bancos poderão deter um nível de capital mais elevado do que o que advém da solução do seu problema de otimização estático.

Assim, assume-se que os bancos resolvem um problema de minimização dos custos relacionados

com o trabalho, o financiamento e o capital ( $w_L L + w_F F + w_K K$ ) sujeito à produção de um nível predeterminado de empréstimos ( $\bar{y}_1$ ) e de outros ativos que rendem juros ( $\bar{y}_2$ ) e à manutenção de um dado nível de capitais próprios ( $e_0$ ):

$$C(y_1, y_2, w_L, w_F, w_K, e) = \min_{L, F, K} (w_L L + w_F F + w_K K) \quad (1)$$

$$\begin{aligned} & \text{s.t.} \\ & F(x, e) \geq \bar{y} \\ & e \geq e_0 \end{aligned}$$

Em que as variáveis são definidas da seguinte forma:

$$C \equiv \sum_k w_k x_k \quad (2)$$

- $y_1$ : Empréstimos concedidos
- $y_2$ : Outros ativos que rendem juros
- $w_L$ : Preço do Trabalho
- $w_F$ : Preço do Financiamento
- $w_K$ : Preço do Capital
- $L$ : Trabalho
- $F$ : Financiamento
- $K$ : Capital
- $e$ : Capitais Próprios

O preço do financiamento é calculado como o rácio entre o fluxo de juros pagos e o valor dos passivos que pagam juros, ao passo que o preço do trabalho é definido como o rácio entre os custos com pessoal e o número de empregados e o preço do capital foi aproximado pelo rácio entre os custos operacionais (excluindo mão-de-obra) e o nível de ativos tangíveis e intangíveis.

É um facto amplamente divulgado que os bancos, tal como as outras empresas, não se comportam estritamente como maximizadores de lucro. Assim, quer devido a problemas de agência, quer devido a diferenças na capacidade dos gestores, alguns bancos estarão mais próximos do comportamento ótimo do que outros. Adicionalmente, tal como é habitual em aplicações empíricas, assume-se que o desempenho de cada banco poderá ser afetado por fatores aleatórios, ao mesmo tempo que as variáveis utilizadas no processo de estimação poderão estar afetadas por erros de medida. Assim, de forma a analisar a eficiência-custo dos bancos portugueses desde o início da década de noventa, a função custo que advém da solução do problema formulado na Equação (1) é estimada recorrendo a modelos de Análise de Fronteiras Estocásticas (SFA). A equação principal a estimar pode ser escrita como:

$$\ln C_{i,t} = \ln C(t_t, y_{r,i,t}, w_{k,i,t}, e_{i,t}) + (v_{i,t} + u_{i,t}) \quad (3)$$

Em que  $C(t_t, y_{r,i,t}, w_{k,i,t}, e_{i,t})$  representa a fronteira de custo estimada e  $C_{i,t}$  os custos incorridos na realidade pelos bancos, pelo que estes têm como limite inferior a soma da fronteira estima-

da com o erro aleatório  $(v_{i,t})$ , que se assume seguir uma distribuição i.i.d.  $N(0, \sigma_v^2)$  e engloba o efeito de eventuais erros de medida e de outros choques aleatórios não controláveis. A soma de  $\ln C(t, y_{r,i,t}, w_{k,i,t}, e_{i,t})$  e  $v_{i,t}$  constitui a fronteira estocástica e  $u_{i,t}$  representa uma variável aleatória não negativa através da qual pode ser medida a ineficiência-custo como a diferença entre o custo observado e a fronteira de custo estocástica. Estão documentados na literatura modelos que diferem nas hipóteses acerca da distribuição de  $u_{i,t}$ . O modelo apresentado em Battese e Coelli (1995) assume que  $u_{i,t}$  é independentemente distribuído como uma variável truncada em zero e segue uma distribuição  $N(m_{i,t}, \sigma_u^2)$ , em que  $m_{i,t} = z_{i,t}\delta$  e, por sua vez,  $z_{i,t}$  é um vetor de variáveis específicas a cada banco e período de tempo e  $\delta$  é um vetor de coeficientes a estimar. Esta especificação tem a vantagem de permitir a identificação do impacto de algumas variáveis (as incluídas no vetor  $z_{i,t}$ ) na eficiência-custo estimada. Neste artigo, as variáveis selecionadas para o efeito foram o rácio entre os empréstimos que entraram em incumprimento há menos de um ano e os empréstimos concedidos (*INC*) como uma medida de risco de crédito, a quota de mercado em crédito (*QM*) como uma medida de dimensão relativa dos bancos, a rentabilidade do capital e do ativo (*ROE* e *ROA*, respetivamente), o rácio entre os capitais próprios e o ativo (*C/A*) e uma medida de liquidez definida como o rácio entre os ativos voláteis e os passivos voláteis (*Rácio Liq.*).

Um modelo alternativo, proposto em Battese e Coelli (1992) especifica  $u_{i,t}$  da seguinte forma:

$$u_{i,t} = u_i \exp(-\eta(t - T)) \quad (4)$$

Em que  $u_i$  é independentemente distribuído como uma variável truncada em zero e segue uma distribuição  $N(\mu, \sigma_u^2)$  sendo que  $\mu$  e  $\eta$  são parâmetros a estimar. Nesta especificação, a ineficiência estimada é específica a cada banco e pode variar ao longo do tempo apesar de, ao contrário do que se verifica no modelo de Battese e Coelli (1995), a ordem dos bancos por grau de eficiência ser constante ao longo do tempo. Caso a estimativa para  $\eta$  não seja estatisticamente significativa, esta restrição pode ser incluída no modelo de forma a maximizar os graus de liberdade ao evitar estimar mais parâmetros do que os necessários. A eficiência-custo do banco  $i$  no período  $t$  é definida como:

$$CE_{i,t} = \frac{C(t, y_{r,i,t}, w_{k,i,t}, e_{i,t})}{C_{i,t}} \in (0,1] \quad (5)$$

O custo observado de um banco sem ineficiência encontra-se sobre a fronteira, de forma a que a sua eficiência é de 100%, ao passo que um banco com uma eficiência de X% apresenta um custo superior à fronteira, pelo que, em teoria, poderia manter o seu nível de atividade com apenas X% dos seus custos.

A escolha da forma funcional utilizada para especificar a função custo deve atingir um equilíbrio entre a flexibilidade e a parcimónia, de forma a obter uma boa aproximação à verdadeira função custo, preservando ao mesmo tempo os graus de liberdade disponíveis e evitando problemas de multicolinearidade. A especificação Cobb-Douglas é em geral considerada demasiado restritiva, ao

passo que a forma funcional *translog* constitui uma aproximação local flexível e a especificação Fourier permite uma aproximação global. O estudo de Berger e Mester (1997) encontrou diferenças estatisticamente mas não economicamente relevantes entre estas duas formas funcionais. Assim, e tendo em conta a dimensão relativamente reduzida da amostra disponível, optou-se neste artigo por estimar a função custo recorrendo a uma forma funcional *translog*, que pode ser representada da seguinte forma:

$$\begin{aligned} \ln C_{i,t} = & \gamma_0 + \gamma_1 t_t + \frac{1}{2} \gamma_{11} t_t^2 + \sum_r \gamma_{1,r} t_t \ln y_{r,i,t} + \sum_k \gamma_{1,k} t_t \ln w_{k,i,t} + \sum_r \gamma_r \ln y_{r,i,t} + \\ & + \sum_k \gamma_k \ln w_{k,i,t} + \frac{1}{2} \left( \sum_r \sum_s \gamma_{r,s} \ln y_{r,i,t} \ln y_{s,i,t} + \sum_k \sum_l \gamma_{k,l} \ln w_{k,i,t} \ln w_{l,i,t} \right) + \\ & + \sum_k \sum_r \gamma_{k,r} \ln w_{k,i,t} \ln y_{r,i,t} + \gamma_e e_{i,t} + \frac{1}{2} \gamma_{ee} e_{i,t}^2 + \\ & + \sum_k \gamma_{e,k} e_{i,t} \ln w_{k,i,t} + \sum_r \gamma_{e,r} \ln e_{i,t} \ln y_{r,i,t} + v_{i,t} + u_{i,t} \end{aligned} \quad (6)$$

Em que são incluídas as habituais restrições teóricas que advêm da teoria da dualidade (simetria e homogeneidade linear nos preços dos *inputs*). As equações para a proporção do custo total correspondente a cada *input* derivadas do lema de Shephard não foram impostas, uma vez que a relação em causa apenas se verifica sob a hipótese de inexistência de ineficiência alocativa. Assim, a medida de ineficiência-X estimada inclui ineficiência técnica bem como ineficiência alocativa.

A base de dados utilizada neste artigo foi construída através das demonstrações financeiras reportadas pelos bancos ao Banco de Portugal e constitui um painel não equilibrado de dados anuais para todos os bancos com atividade em Portugal entre 1991 e 2004<sup>5</sup>. Os empréstimos concedidos foram ajustados para operações de titularização, uma vez que o banco que origina um empréstimo é em geral responsável pelo seu acompanhamento mesmo que o referido empréstimo seja titularizado. Assim, caso esta correção não fosse feita, a eficiência dos bancos envolvidos em operações de titularização seria subestimada<sup>6</sup>. Os empréstimos em incumprimento não foram incluídos na definição de *output* uma vez que, no essencial, constituem uma rubrica do balanço que não produz proveitos. Assim, este procedimento tem em conta os diferentes níveis de risco de crédito na carteira dos bancos e corrige implicitamente para diferenças no nível e na qualidade das atividades de análise e acompanhamento do risco de crédito dos clientes. É, contudo, importante ressaltar que este argumento assume implicitamente que a carteira de crédito dos bancos é homogénea, ao passo que, na prática, cada banco pode operar em segmentos de mercado com diferentes níveis de risco de crédito. Apesar de terem sido encetados alguns esforços na seleção de uma amostra de bancos com uma estrutura de atividade semelhante, uma solução completa para este problema passaria pela definição de um *output* para cada categoria de empréstimos. No caso concreto deste artigo, tal solução não seria viável, uma vez que a amostra disponível não permite acomodar o aumento no

(5) A razão pela qual não foram utilizados dados mais recentes prende-se com alterações que se verificaram em 2005 no regime contabilístico que levam a uma quebra de série nos dados.

(6) As operações de titularização em Portugal começaram a realizar-se em 1997 e a sua importância cresceu significativamente no período mais recente, sendo que em 2004 o volume de empréstimos titularizados correspondia a cerca de 6% do total dos empréstimos concedidos. Contudo, é importante destacar a presença de alguma heterogeneidade no que respeita a esta proporção, sendo que um banco em particular apresentava em 2004 um peso de empréstimos titularizados no total de empréstimos concedidos de 34% em base não consolidada.

número de parâmetros a estimar que esta abordagem implicaria.

Apesar do reporte de demonstrações financeiras ao Banco de Portugal ser obrigatório para todos os bancos a operar em Portugal, apenas foram incluídos na análise os bancos de retalho universais. Esta opção prende-se com o objetivo de garantir que a análise se concentra em bancos cuja tecnologia é suficientemente semelhante para que seja legítimo incluí-los na mesma função custo. Adicionalmente, de forma a evitar distorções provocadas pela relação atípica entre os custos e o *output* dos bancos que tende a verificar-se na fase inicial da sua atividade, os bancos mais recentes só foram incluídos na análise desde o seu terceiro ano de funcionamento.

Da aplicação destes critérios de seleção resultou uma amostra de 25 bancos, constituindo um total 254 pares banco-ano. Para cada um dos anos considerados, a amostra cobre um mínimo de 77% do total dos empréstimos, 80% do total dos ativos e 87% do total dos depósitos do sistema bancário português. Adicionalmente, a quota de mercado dos cinco maiores bancos, quando medida em termos do total de ativos da amostra, aumentou de cerca de 57% em 1992 para um valor próximo de 70% em 2004, evolução semelhante à observada para o total do sistema bancário.

### 3. RESULTADOS EMPÍRICOS

O Quadro 1 apresenta os principais resultados da análise levada a cabo com base na estimação da Equação (6). Uma vez que os dados foram expressos como desvios em relação à média, as principais elasticidades da função custo podem ser lidas diretamente através de parâmetros isolados<sup>7</sup>. Assim, e de forma a clarificar a exposição, a maioria dos termos cruzados não são apresentados no Quadro 1<sup>8</sup>. Uma análise preliminar dos resultados da estimação indica que a elasticidade do custo em relação ao preço dos diferentes *inputs* é positiva. Adicionalmente, o preço do *input* em relação ao qual os custos apresentam uma reação mais forte é o preço do financiamento, resultado que seria de esperar tendo em conta que os custos de financiamento constituem a proporção mais significativa do custo total. O valor da soma dos parâmetros relativos aos *outputs* é próximo de um, indicando a presença de resultados próximos de constantes à escala na média da amostra. O facto de o parâmetro correspondente ao termo de interação entre os dois *outputs* ser negativo indica a existência de economias de gama na produção conjunta de empréstimos e de outros ativos que rendem juros. Existe evidência da presença de progresso tecnológico avaliado na média da amostra e, como seria de esperar, os bancos com um nível de capitais próprios mais elevado tendem a apresentar menores custos relacionados com os outros *inputs*. O facto da estimativa obtida para o parâmetro  $\gamma$  ser estatisticamente significativa nas diferentes especificações indica que a estimação da função custo como uma fronteira é apropriada, uma vez que este parâmetro mede a proporção da variância total do erro de estimação que se deve a ineficiência-X e não ao clássico erro de estimação aleatório. Adicionalmente, a estimativa obtida para  $\gamma$  é mais reduzida nos modelos que incluem variáveis explicativas como determinantes da ineficiência, resultado que seria de esperar uma vez que as variáveis incluídas no vetor  $z_{i,t}$  capturam parte da variância de  $u_{i,t}$ .

(7) Para uma discussão sobre este aspeto, ver Boucinha, Ribeiro e Weyman-Jones (2009).

(8) Os resultados completos das estimações estão disponíveis sob pedido.

Quadro 1

RESULTADOS DA ESTIMAÇÃO DE FRONTEIRAS DE CUSTOS							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\ln(W_L)$	0.1260 <i>0.02</i>	0.1253 <i>0.02</i>	0.1266 <i>0.02</i>	0.1278 <i>0.02</i>	0.1519 <i>0.02</i>	0.1148 <i>0.02</i>	0.1147 <i>0.02</i>
$\ln(W_F)$	0.6233 <i>0.01</i>	0.6243 <i>0.01</i>	0.6226 <i>0.01</i>	0.6213 <i>0.01</i>	0.6019 <i>0.01</i>	0.6364 <i>0.01</i>	0.6367 <i>0.01</i>
$\ln(y_1)$	0.5800 <i>0.01</i>	0.5798 <i>0.01</i>	0.5798 <i>0.01</i>	0.6040 <i>0.01</i>	0.5872 <i>0.01</i>	0.5562 <i>0.01</i>	0.5566 <i>0.01</i>
$\ln(y_2)$	0.4472 <i>0.01</i>	0.4468 <i>0.01</i>	0.4475 <i>0.01</i>	0.4679 <i>0.01</i>	0.4409 <i>0.01</i>	0.4589 <i>0.01</i>	0.4591 <i>0.01</i>
$\ln(y_1)*\ln(y_2)$	-0.2201 <i>0.02</i>	-0.2208 <i>0.02</i>	-0.2194 <i>0.02</i>	-0.2417 <i>0.01</i>	-0.2079 <i>0.02</i>	-0.2496 <i>0.02</i>	-0.2499 <i>0.01</i>
$t$	-0.0192 <i>0.00</i>	-0.0191 <i>0.00</i>	-0.0194 <i>0.00</i>	-0.0195 <i>0.00</i>	-0.0236 <i>0.00</i>	-0.0142 <i>0.00</i>	-0.0143 <i>0.00</i>
$\ln(e)$	-0.0061 <i>0.02</i>	-0.0053 <i>0.02</i>	-0.0067 <i>0.02</i>	-0.0539 <i>0.01</i>	-0.0002 <i>0.02</i>	-0.0486 <i>0.01</i>	-0.0492 <i>0.01</i>
$INC$	0.8847 <i>0.12</i>	0.8842 <i>0.12</i>	0.8845 <i>0.12</i>	0.8978 <i>0.12</i>	0.6681 <i>0.12</i>		
$QM$	-1.0852 <i>0.30</i>	-1.0759 <i>0.29</i>	-1.0835 <i>0.30</i>	-0.9847 <i>0.31</i>	-1.1822 <i>0.40</i>		
$C/A$	-0.7913 <i>0.27</i>	-0.7739 <i>0.27</i>	-0.7865 <i>0.27</i>		-0.7525 <i>0.32</i>		
$ROE$	-0.0103 <i>0.03</i>						
$ROA$		-0.40477 <i>0.61</i>					
$R\acute{a}cio\ Liq.$					0.01 <i>0.00</i>		
$\mu$						0.0682 <i>0.06</i>	0.0713 <i>0.04</i>
$\eta$						0.0021 <i>0.03</i>	
$\gamma$	0.7126 <i>0.16</i>	0.7169 <i>0.16</i>	0.7130 <i>0.16</i>	0.7033 <i>0.18</i>	0.5442 <i>0.31</i>	0.77 <i>0.1295</i>	0.77 <i>0.1258</i>

Fonte: Cálculos dos autores.

Nota: Os desvios-padrão são apresentados em itálico. A constante e a maioria dos termos cruzados da especificação da função custo foram omitidos por não terem uma interpretação direta. Os resultados completos das estimações estão disponíveis sob pedido.

As primeiras cinco colunas do quadro apresentam os resultados da estimação de modelos em que a ineficiência é definida em função de uma série de características dos bancos. Assim, conclui-se que os bancos com um rácio mais elevado entre o fluxo de empréstimos em incumprimento e os empréstimos concedidos tendem a ter níveis mais elevados de ineficiência-X. Este resultado sugere que o grau de incumprimento tende a ser representativo da qualidade da gestão do banco, *i.e.* os gestores com piores resultados na avaliação e gestão do risco de crédito tendem também a ter um desempenho comparativamente pior a nível do controlo dos custos. Os bancos com uma quota de mercado mais elevada em termos de empréstimos tendem a ser mais eficientes, o que sugere que os bancos de maior dimensão conseguem atrair gestores mais competentes. Esta hipótese poderá, por sua

vez, refletir salários mais elevados e outros benefícios que os bancos mais pequenos não possam oferecer à sua gestão de topo, bem como o prestígio associado à liderança dos maiores bancos. Os bancos com menor alavancagem tendem a apresentar valores de eficiência custo mais elevados, o que poderá estar relacionado com a presença de menores custos de agência. A rentabilidade não se mostrou como um determinante relevante da eficiência-X. Por outro lado, o rácio de liquidez dos bancos mostrou-se positivamente correlacionado com o nível de ineficiência, o que sugere que os bancos com uma gestão da liquidez menos agressiva tenderão a também a ser menos eficientes em termos de custos. É, contudo, importante salientar que os bancos que optam por deter menores reservas de liquidez estarão sujeitos a um risco de liquidez mais elevado, que poderá não se refletir de forma adequada nos preços estabelecidos nos mercados por grosso em períodos de funcionamento normal dos mercados, situação que prevaleceu durante o período em análise.

Na sétima coluna do Quadro 1, a estimativa para a ineficiência-X é definida em função do tempo, tal como descrito na Equação (4). Contudo, uma vez que a estimativa para o parâmetro  $\eta$  não se mostrou estatisticamente significativa, a coluna seguinte apresenta os resultados da estimação de um modelo no qual se impõe a restrição do valor deste parâmetro ser zero, evitando assim a perda de graus de liberdade associada à estimação de um parâmetro redundante. Esta é a especificação utilizada para a análise cujos resultados são discutidos nas próximas sub-secções, a primeira das quais se refere às estimativas obtidas para os custos marginais dos bancos, incluindo a sua evolução ao longo do tempo. A subsecção seguinte dedica-se às estimativas para o custo sombra dos capitais próprios dos bancos ao passo que na terceira subsecção são discutidos os resultados relativos à eficiência de escala. As subsecção que se seguem apresentam resultados relativos à eficiência-X e ao progresso tecnológico. A última subsecção dedica-se à análise da produtividade total de fatores do sistema bancário, nomeadamente à sua quantificação, evolução e decomposição de forma a avaliar o contributo da evolução da tecnologia, da aproximação dos bancos às melhores práticas e de movimentos dos bancos para pontos diferentes da mesma função custo.

### 3.1. Custo marginal

Recorrendo aos parâmetros estimados para a função custo, podem ser obtidas estimativas para o custo marginal na produção de cada *output* através da seguinte expressão:

$$mc_{r,i,t} \equiv \frac{\partial C_{i,t}}{\partial y_{r,y,t}} = \frac{C_{i,t}}{y_{r,i,t}} \frac{\partial \ln C_{i,t}}{\partial \ln y_{r,i,t}} \quad (7)$$

A Equação (7) permite obter estimativas para os custos marginais de cada banco, tanto para a produção de empréstimos como de outros ativos que rendem juros. Assim, as séries temporais apresentadas no Quadro 2 foram construídas através da agregação das estimativas individuais, utilizando como pesos a quota de mercado de cada banco no mercado de empréstimos. O facto das estimativas para os custos marginais terem diminuído significativamente ao longo do tempo seria de esperar tendo em conta que os custos de financiamento constituem uma parte substancial dos custos variáveis dos bancos e as taxas de juro exibiram uma forte tendência decrescente durante

o período em análise (Gráfico 1). Assim, é também interessante investigar a evolução do custo marginal associado aos recursos não financeiros dos bancos no mesmo período. É possível obter uma medida para esta variável deduzindo o custo unitário do financiamento de cada banco ao custo marginal estimado. Tal como ilustrado nas colunas 5 e 6 do Quadro 2 e no Gráfico 2, esta medida apresenta igualmente uma tendência decrescente ao longo do tempo sugerindo que, apesar de ter contribuído para o perfil exibido pelos custos marginais ao longo do tempo, o comportamento das taxas de juro por si só é insuficiente para o explicar. Durante o período em análise verificou-se uma progressiva alteração na estrutura da carteira de empréstimos dos bancos portugueses, com um aumento da importância dos empréstimos a particulares em contrapartida de uma diminuição do peso dos empréstimos ao setor público. Esta alteração de estrutura deverá ter contribuído para um aumento do custo marginal dos empréstimos concedidos. Como tal, a significativa redução observada nas estimativas para o custo marginal operacional dos empréstimos não deverá ter sido motivada por alterações na composição da carteira de empréstimos dos bancos.

Tal como ilustrado no Gráfico 2 e documentado no Quadro 2, o custo marginal dos empréstimos tem em geral apresentado um valor superior ao correspondente aos outros ativos que rendem juros, indicando que conceder um empréstimo adicional consome mais recursos do que investir em títulos, resultado que deverá estar relacionado com os custos de avaliação e acompanhamento associados à concessão de empréstimos. Contudo, esta diferença tem-se tornado menos relevante ao longo do tempo. De forma a interpretar esta trajetória, é importante recordar que o *output* definido como outros ativos que rendem juros inclui produtos bastante heterogéneos. De facto, enquanto durante o início da década de 90 os bancos detinham um volume considerável de títulos do tesouro e de depósitos junto do banco central, com a liberalização dos sistema bancário e a inovação financeira, a carteira de investimentos financeiros dos bancos passou a incluir ativos mais sofisticados que, devido à sua maior complexidade, exigem uma maior utilização de recursos.

**Quadro 2**

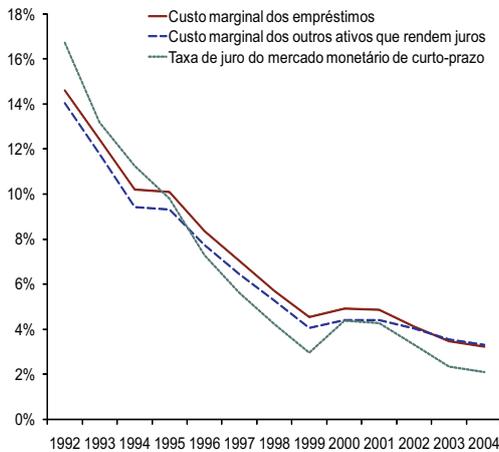
ESTIMATIVAS PARA OS CUSTOS MARGINAIS AVALIADAS NA MÉDIA PONDERADA DA AMOSTRA								
Ano	Custo marginal dos empréstimos	Custo marginal dos outros ativos que rendem juros	Taxa de juro do mercado monetário de curto-prazo	Preço implícito do financiamento	Custo marginal não financeiro dos empréstimos	Custo marginal não financeiro dos outros ativos que rendem juros	Taxa de juro implícita dos empréstimos	Margem nos empréstimos
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1992	14.59%	14.03%	16.72%	10.74%	3.85%	3.29%	17.15%	2.56%
1993	12.43%	11.79%	13.17%	8.94%	3.48%	2.84%	15.45%	3.02%
1994	10.21%	9.41%	11.23%	7.01%	3.20%	2.40%	13.08%	2.86%
1995	10.10%	9.31%	9.79%	7.09%	3.01%	2.21%	12.26%	2.15%
1996	8.35%	7.73%	7.27%	5.71%	2.64%	2.02%	10.72%	2.37%
1997	7.04%	6.45%	5.61%	4.65%	2.39%	1.80%	9.24%	2.20%
1998	5.69%	5.26%	4.23%	3.59%	2.10%	1.66%	7.44%	1.75%
1999	4.53%	4.05%	2.96%	2.65%	1.89%	1.41%	5.78%	1.25%
2000	4.93%	4.40%	4.39%	3.28%	1.64%	1.11%	6.13%	1.20%
2001	4.85%	4.41%	4.26%	3.46%	1.40%	0.95%	6.30%	1.45%
2002	4.12%	4.03%	3.32%	2.76%	1.35%	1.26%	5.36%	1.24%
2003	3.46%	3.54%	2.33%	2.28%	1.17%	1.26%	4.56%	1.10%
2004	3.21%	3.30%	2.11%	2.00%	1.21%	1.29%	4.21%	1.00%

**Fontes:** Banco de Portugal e cálculos dos autores.

**Nota:** No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização.

Gráfico 1

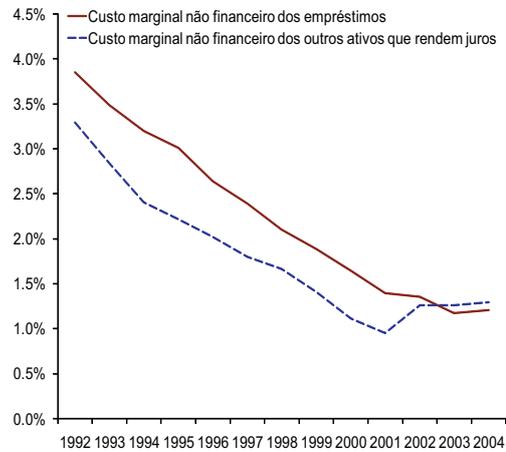
## CUSTOS MARGINAIS



Fontes: Banco de Portugal e cálculos dos autores.

Gráfico 2

## CUSTOS MARGINAIS NÃO FINANCEIROS



Fonte: Cálculos dos autores.

Tal como ilustrado na Coluna 7 do Quadro 2, é ainda possível calcular uma taxa de juro implícita para os empréstimos concedidos por cada banco, através de dados relativos às receitas relacionadas com juros de empréstimos concedidos pelos bancos e ao valor dos referidos empréstimos. Subtraindo a esta taxa de juro o custo marginal relevante, obtém-se uma medida da margem entre o preço praticado pelos bancos e o seu custo marginal. De acordo com os resultados apresentados na última coluna do Quadro 2 e no Gráfico 3, esta medida tem diminuído ao longo do tempo, o que é consistente com o resultado encontrado em Boucinha e Ribeiro (2009), segundo o qual a concorrência no sistema bancário terá aumentado durante o período em análise<sup>9</sup>.

### 3.2. Custo sombra dos capitais próprios

Uma vez que a função custo estimada inclui o nível dos capitais próprios como um *input* fixo, permite o cálculo de uma medida do custo sombra do capital de acordo com a seguinte expressão:

$$w_k^* = -\frac{\partial C_{i,t}}{\partial e_{i,t}} = -\frac{C_{i,t}}{e_{i,t}} \frac{\partial \ln C_{i,t}}{\partial \ln e_{i,t}} \quad (8)$$

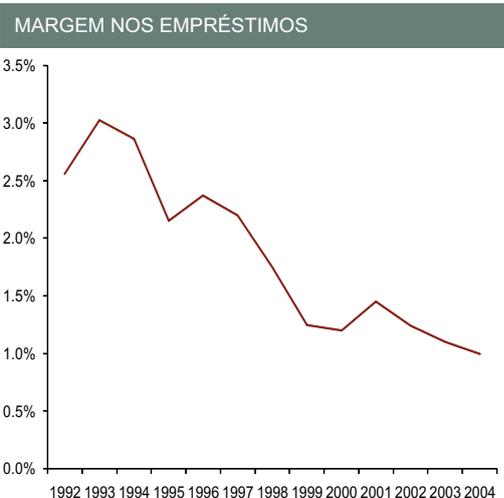
O objetivo de calcular o custo sombra dos capitais próprios é obter uma medida da disponibilidade dos bancos a pagar pelo capital, uma vez que indica o montante que estes poupariam nos restantes custos como resultado de um reforço da sua capitalização<sup>10</sup>.

Tal como ilustrado no Gráfico 4, a série temporal obtida através da agregação das estimativas para

(9) A medida de taxa de juro implícita utilizada foi calculada com base em receitas de juros e em valores de empréstimos concedidos que não incluem os empréstimos em incumprimento. Assim, esta medida aproxima a taxa de juro que os bancos cobram aos seus clientes que, devido à ocorrência de episódios de incumprimento, deverá ser superior à que os bancos recebem de facto. Assim, a diminuição verificada no grau de incumprimento ao longo do período da amostra deverá também ter contribuído para a diminuição observada na margem dos bancos. Ainda assim, construindo uma taxa de juro que constitui um limite inferior para aquela que os bancos recebem de facto, ao considerar o valor dos empréstimos em incumprimento mas não a receita de juros correspondente, mantém-se a diminuição observada na margem associada aos empréstimos concedidos. Assim, este comportamento não estará a ser apenas motivado pela diminuição do incumprimento observada ao longo do período da amostra.

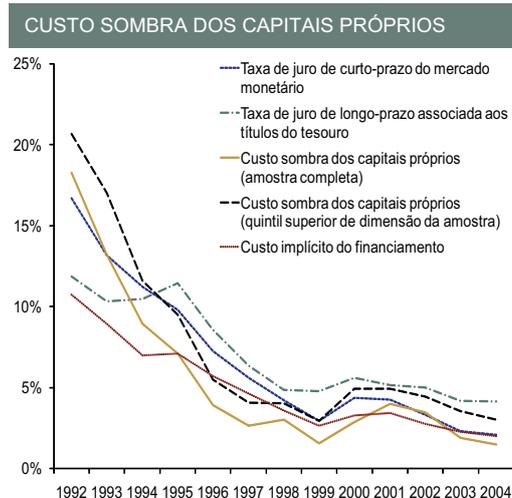
(10) A este respeito, há no entanto que ter em conta as limitações do modelo utilizado ao ser desenvolvido no contexto de um modelo de otimização estático e recorrer a dados contabilísticos.

Gráfico 3



Fonte: Cálculos dos autores.

Gráfico 4



Fontes: Banco de Portugal e cálculos dos autores.

o custo sombra dos capitais próprios é fortemente correlacionada com as taxas de juro de mercado e com o custo médio ponderado de financiamento dos bancos através de depósitos e dos mercados de financiamento por grosso. Este resultado é consistente com a interpretação do capital próprio como uma fonte de financiamento dos bancos em si mesma, pelo que os custos de financiamento deverão ser os mais afetados pelo valor dos capitais próprios.

O valor obtido para a variável em análise, apresentado no Quadro 3, é inferior àquele que é geralmente considerado como um valor razoável para o verdadeiro preço do capital. Este facto não é surpreendente e sugere que a opção de considerar os capitais próprios como um *input* fixo ao invés de um *input* variável está correta, uma vez que aponta para um efeito relevante das restrições de ordem regulamentar e reputacionais na escolha do nível de capitalização do bancos, de forma que os bancos tendem na prática a deter um nível de capital superior ao que estaria subjacente à solução do seu problema de otimização estático.

Com o objetivo de investigar quais os fatores que determinam as diferenças entre o custo sombra dos capitais próprios dos bancos, foi estimada uma regressão desta variável num conjunto de variáveis específicas a cada banco, incluindo o rácio de capital e a rentabilidade dos capitais e três variáveis dicotómicas que identificam os bancos públicos, as sucursais de instituições de crédito com sede no estrangeiro e os bancos que constituem o quintil superior da distribuição de dimensão em cada ano<sup>11</sup>.

Optou-se por incluir na estimação o valor desfasado do rácio de capital dos bancos e da rentabilidade dos capitais de forma a evitar problemas de simultaneidade entre estas variáveis e o custo sombra dos capitais próprios. Uma vez que a variável binária que identifica as sucursais de instituições de crédito com sede no estrangeiro é constante ao longo do tempo, a identificação do coeficiente as-

(11) Uma vez que a variável dependente desta regressão é já uma estimativa, os desvios-padrão que resultam da mesma não são válidos, uma vez que não têm em conta a variância da referida estimativa.

## Quadro 3

CUSTO SOMBRA DOS CAPITALS PRÓPRIOS						
Ano	Custo sombra dos capitais próprios (amostra completa)	Custo sombra dos capitais próprios (quartil superior de dimensão da amostra)	Taxa de juro de curto-prazo do mercado monetário	Taxa de juro de longo-prazo associada aos títulos do tesouro	Rácio capitais próprios/ativo	Custo implícito do financiamento
	(1)	(2)	(3)	(4)	(5)	(6)
1992	18.26%	20.68%	16.72%	11.86%	7.56%	10.74%
1993	13.20%	17.02%	13.17%	10.33%	7.41%	8.94%
1994	8.94%	11.60%	11.23%	10.48%	6.97%	7.01%
1995	7.13%	9.52%	9.79%	11.47%	6.54%	7.09%
1996	3.93%	5.51%	7.27%	8.56%	6.29%	5.71%
1997	2.67%	4.09%	5.61%	6.36%	6.07%	4.65%
1998	3.04%	4.04%	4.23%	4.88%	6.39%	3.59%
1999	1.55%	2.96%	2.96%	4.78%	6.50%	2.65%
2000	2.87%	4.92%	4.39%	5.60%	6.04%	3.28%
2001	4.01%	4.94%	4.26%	5.16%	5.85%	3.46%
2002	3.46%	4.46%	3.32%	5.01%	6.13%	2.76%
2003	1.90%	3.56%	2.33%	4.18%	6.36%	2.28%
2004	1.51%	3.04%	2.11%	4.14%	6.33%	2.00%

Fontes: Banco de Portugal e cálculos dos autores.

Nota: No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização.

sociado a esta variável não é possível numa regressão que inclua efeitos fixos para cada banco. Por este motivo são apresentadas regressões estimadas assumindo efeitos fixos e efeitos aleatórios.

Os resultados destas estimações, apresentados no Quadro 4, sugerem que os bancos mais capitalizados tendem a deter um maior custo sombra dos capitais próprios. A rentabilidade surge também com um efeito positivo, o que poderá em parte refletir níveis de risco mais elevados. Por outro lado, possivelmente associado a um menor risco apercebido pelos mercados na respetiva dívida, os bancos de capitais públicos tendem a apresentar um custo sombra dos capitais mais reduzido. As sucursais de instituições financeiras com sede no estrangeiro representam em geral uma proporção reduzida dos ativos do grupo bancário a que pertencem, exercendo assim uma influência negligenciável sobre o *rating* de crédito do grupo e obtêm frequentemente financiamento diretamente junto da sua sede. Neste contexto, o facto de estes bancos apresentarem em média valores inferiores para o custo sombra dos capitais próprios surge naturalmente. Por outro lado, os bancos de maior dimensão, que tendem a ser caracterizados por uma maior transparência, e cujo capital tem maior probabilidade de ser cotado nos mercados acionistas, tendem a ter um custo sombra dos capitais próprios mais elevado.

**Quadro 4**

DETERMINANTES DO CUSTO SOMBRA DOS CAPITAIS PRÓPRIOS		
	Modelo de efeitos aleatórios	Modelo de efeitos fixos
	(1)	(2)
$(\text{Capitais Próprios/Ativo})_{i,t-1}$	0.84 <i>0.14</i>	0.71 <i>0.14</i>
$ROE_{i,t-1}$	0.14 <i>0.03</i>	0.13 <i>0.03</i>
Variável binária que identifica bancos de capitais públicos	-0.05 <i>0.02</i>	-0.04 <i>0.02</i>
Variável binária que identifica sucursais de instituições de crédito com sede no estrangeiro	-0.08 <i>0.04</i>	
Variável binária que identifica os bancos do quintil superior de dimensão da amostra	0.02 <i>0.01</i>	0.04 <i>0.02</i>
Teste de Hausman ( <i>p-value</i> )		0.995

**Fonte:** Cálculos dos autores.

**Notas:** Os desvios-padrão são apresentados em itálico. A estimação inclui efeitos fixos para o tempo e uma constante.

### 3.3. Eficiência de escala

A investigação da existência de economias de escala tem sido alvo de extensa discussão na literatura. Apesar de haver argumentos teóricos que apoiam a sua existência e de esta ser frequentemente alegada pela gestão dos bancos como uma motivação para a participação em fusões e aquisições, são frequentes os estudos empíricos que não verificam a sua presença nos dados. A identificação das economias de escala tem associadas implicações relevantes, uma vez que permite inferir sobre a desejabilidade de uma determinada estrutura de mercado de um ponto de vista tecnológico.

Esta secção dedica-se à avaliação da presença de economias de escala desde a liberalização do sistema bancário português. Uma medida de economias de escala é tipicamente obtida de acordo com a seguinte expressão:

$$SE_{i,t} \equiv \sum_r \frac{\partial \ln C_{i,t}}{\partial \ln y_{r,y,t}} \quad (9)$$

Uma elasticidade do custo total em relação ao valor dos empréstimos inferior (superior) a um indica a presença de economias (deseconomias) de escala. Tal como ilustrado no Quadro 5, foram identificadas deseconomias de escala (de acordo com a definição acima) para o início da década de 90 de forma que, tudo o resto constante, um aumento na dimensão de um banco levaria a um aumento mais do que proporcional dos custos. Durante o período mais recente, a estimativa obtida para o parâmetro de escala foi ligeiramente inferior a um, embora não estatisticamente diferente deste valor, indicando a presença de rendimentos à escala virtualmente constantes. Há, contudo, que ter em conta que a elasticidade calculada de acordo com a Equação (9) constitui uma medida de economias de escala de curto prazo, no sentido em que não se permitem variações no nível dos capitais próprios dos bancos. Adicionalmente, uma vez que a definição dos custos utilizada não inclui o custo dos capitais próprios, a medida de economias de escala acima apresentada constitui na prática uma medida em termos de fluxos de caixa. Assim, esta deverá sobrestimar o verdadeiro parâmetro de

Quadro 5

ECONOMIAS DE ESCALA				
Ano	Economias de escala em termos de fluxos de caixa (SE)	p-value (H0: SE=1)	Economias de escala em termos económicos (ESE)	p-value (H0: ESE=1)
1992	1.0919	0.00	0.9698	0.01
1993	1.0747	0.00	0.9678	0.01
1994	1.0505	0.02	0.9664	0.00
1995	1.0349	0.08	0.9687	0.01
1996	1.0156	0.41	0.9665	0.00
1997	0.9995	0.98	0.9628	0.00
1998	1.0010	0.96	0.9533	0.00
1999	0.9882	0.57	0.9457	0.00
2000	0.9906	0.66	0.9431	0.00
2001	0.9933	0.79	0.9409	0.00
2002	0.9909	0.71	0.9363	0.00
2003	0.9770	0.35	0.9347	0.00
2004	0.9658	0.20	0.9297	0.00
1992-2004	0.9964	0.86	0.9447	0.00

Fonte: Cálculos dos autores.

Notas: No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização. SE denota as economias de escala definidas de acordo com a Equação (9) e ESE refere-se à definição de economias de escala apresentada na Equação (14).

escala, uma vez que o facto do nível de capital ser mantido fixo implica que qualquer aumento da atividade terá de ser financiado inteiramente por dívida sujeita ao pagamento de juros, pelo que são assumidos aumentos para o custo da dívida superiores ao que seria realista.

Em teoria seria possível obter uma medida de economias de escala que permita que o nível de capitais próprios se altere em função de alterações na atividade, estimando uma função custo em que os capitais sejam tratados de forma semelhante aos outros *inputs*. Contudo, tal como discutido acima, esta opção é problemática uma vez que existem restrições importantes à escolha da estrutura de capital dos bancos. Adicionalmente, mesmo no período mais recente, apenas um número reduzido de bancos portugueses têm o seu capital cotado em bolsa, pelo que não há uma forma óbvia de obter uma estimativa para o preço dos capitais próprios.

Em alternativa, tal como destacado por Hughes, Mester e Moon (2001), citando uma proposta original de Hughes (1999), é possível calcular uma medida de economias de escala em termos económicos, e não apenas em termos de fluxos de caixa, assumindo que o nível de capitais próprios observado minimiza os custos para o custo sombra do capital, uma vez que neste caso se verifica a seguinte expressão:

$$C(t, y_r, w_k, w_e^*) = C(t, y_r, w_k, e) + w_e^* e \quad (10)$$

Através da expressão acima, é possível calcular uma medida de economias de escala em termos económicos como:

$$ESE_{i,t} \equiv \sum_r \frac{\partial C(t, y_r, w_k, w_e^*)}{\partial y_r} \frac{y_r}{C(t, y_r, w_k, w_e^*)} \quad (11)$$

Uma vez que o nível de capital  $e$  minimiza o custo económico, o custo marginal restrito é igual ao custo marginal de longo prazo:

$$\frac{\partial C(t, y_r, w_k, w_e^*)}{\partial y_r} = \frac{\partial C(t, y_r, w_k, e)}{\partial y_r} \quad (12)$$

Com base neste resultado e na definição do custo sombra do capital apresentada na Equação (8), a Equação (11) pode ser escrita da seguinte forma:

$$ESE_{i,t} \equiv \sum_r \frac{\partial C(t, y_r, w_k, e)}{\partial y_r} \frac{y_r}{C(t, y_r, w_k, e) - \frac{\partial C}{\partial e} e} \quad (13)$$

ou

$$ESE_{i,t} \equiv \frac{\sum_r \frac{\partial \ln C_{i,t}}{\partial y_{r,i,t}}}{1 - \frac{\partial \ln C_{i,t}}{\partial \ln e_{i,t}}} \quad (14)$$

O Quadro 5 apresenta uma série temporal das estimativas para as economias de escala obtida através da agregação das estimativas individuais resultantes da avaliação da Equação (14). Apesar de esta medida apresentar o mesmo perfil descendente da medida restrita calculada de acordo com a Equação (9), o seu valor para cada ano é consideravelmente inferior. Assim, e tendo em conta que o nível de capitais próprios dos bancos pode variar de acordo com a atividade, são identificadas economias de escala estatisticamente significativas para a totalidade do período em análise, sugerindo que o processo de concentração observado no sistema bancário português terá sido, pelo menos em parte, motivado pela oportunidade de aumentar a eficiência na produção.

Os valores obtidos para as economias de escala são mais expressivos no final da amostra do que no início dos anos 90. Este resultado deverá estar ligado às alterações na tecnologia dos bancos catalisadas pelo progresso tecnológico. De facto, a maior automatização dos serviços deverá ter permitido uma diminuição nos custos variáveis dos bancos, em contrapartida de um investimento inicial significativo em tecnologia relacionada, a título de exemplo, com a armazenagem e processamento de informação e com canais de comunicação. Por sua vez, estes desenvolvimentos tecnológicos permitiram estabelecer uma densa rede de caixas automáticas e de outros pontos distribuição remota de serviços, tais como as páginas dos bancos na *internet*, com as correspondentes poupanças associadas à diminuição da necessidade de mão-de-obra e de balcões. Outro fator que deverá ter contribuído para um aumento das economias de escala no período mais recente prende-se com a crescente internacionalização da atividade bancária motivada pelo progresso tecnológico, pela inovação financeira e pelo aumento da integração económica entre os países membros da União Europeia. De facto, a expansão do mercado relevante para a atividade bancária para além das fronteiras nacionais abriu novas oportunidades de crescimento ao mesmo tempo que, em certa medida, expôs os bancos a concorrência internacional por parte de instituições não residentes. Adicionalmente, mesmo os bancos de maior dimensão dentro do sistema financeiro português têm uma dimensão relativamente reduzida quando comparados com as maiores instituições estrangeiras.

### 3.4. Eficiência-custo

No Quadro 6 são apresentadas as estimativas obtidas para a eficiência-custo dos bancos Portugueses entre 1992 e 2004. Tal como acima referido, os resultados foram obtidos com base numa especificação que não inclui explicitamente variáveis como determinantes da ineficiência. Adicionalmente, a estimativa obtida para  $\eta$  – o parâmetro que mede a variação na eficiência ao longo do tempo – não se mostrou estatisticamente significativa, pelo que este parâmetro foi eliminado da estimação, substituindo-se o seu valor por zero. Assim, não existe evidência de alterações ao longo do período em análise na distância a que os bancos se situam da fronteira que representa as melhores práticas no setor<sup>12</sup>.

A estimativa agregada para a ineficiência-X é próxima de 91%, sugerindo que os bancos portugueses poderiam, em teoria, ter para a mesma atividade incorrido apenas 91% dos custos observados. Importa, contudo, referir que esta estimativa apresenta alguma heterogeneidade, sendo que os valores obtidos variam entre um mínimo de 84% e um máximo de 99%.

#### Quadro 6

ESTIMATIVAS PARA A EFICIÊNCIA-CUSTO	
Ano	Eficiência-X
1992	90.89%
1993	91.11%
1994	91.14%
1995	91.08%
1996	91.02%
1997	91.03%
1998	90.50%
1999	90.53%
2000	90.63%
2001	90.66%
2002	90.67%
2003	90.71%
2004	90.82%
1992-2004	90.76%

Fonte: Cálculos dos autores.

Notas: No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização.

### 3.5. Progresso tecnológico

A função custo estimada inclui uma tendência especificada como termo *translog*, que permite o cálculo de progresso tecnológico neutro segundo a definição de Hicks e não neutro. Uma medida do progresso tecnológico, *i.e.* de deslocamentos da fronteira de custo motivados pela adoção de técnicas de produção mais eficientes, é obtida através do cálculo de  $\frac{\partial \ln C}{\partial t}$ <sup>13</sup>.

(12) O facto do valor agregado da estimativa para a eficiência-custo apresentado no Quadro 6 não ser constante, apesar da estimativa para a eficiência de cada banco não variar ao longo do tempo (tal como se infere pela insignificância estatística do parâmetro  $\eta$ ), deve-se a um efeito de composição. De facto, devido a alterações na quota de mercado dos bancos, os pesos utilizados na agregação (o valor dos empréstimos concedidos) não são constantes e, devido à ocorrência de fusões/aquisições e à emergência de bancos novos, a estimação tem por base um painel de dados não equilibrado.

(13) Note-se que esta medida deverá subestimar o verdadeiro progresso tecnológico caso se verifiquem aumentos na qualidade/variedade dos produtos oferecidos ao longo do tempo.

Tal como ilustrado no Quadro 7, o progresso tecnológico foi muito reduzido durante a primeira metade da década de 90, não sendo possível rejeitar estatisticamente a hipótese de o seu valor ser nulo. Mais recentemente, à medida que os bancos se foram ajustando ao processo de liberalização e se desenvolveram os processos de consolidação e inovação financeira, verificou-se uma intensificação do progresso tecnológico. Estes desenvolvimentos devem ser enquadrados no contexto da integração financeira a um nível global, que catalisou a adoção célere de tecnologias mais eficientes. A estimativa obtida para o progresso tecnológico em 2004 deverá ser interpretada como indicando que, nesse ano, os bancos portugueses que operaram de acordo com as melhores práticas da indústria poderiam obter o mesmo nível de atividade registado no ano anterior com custos 3.2% mais baixos.

**Quadro 7**

PROCESSO TECNOLÓGICO		
Ano	Progresso tecnológico ( <i>eti</i> )	<i>p-value</i> ( $H_0: eti=0$ )
	(1)	(2)
1992	0.19%	0.62
1993	-0.08%	0.42
1994	-0.33%	0.22
1995	-0.67%	0.11
1996	-0.96%	0.02
1997	-1.29%	0.00
1998	-1.44%	0.01
1999	-1.76%	0.03
2000	-2.23%	0.00
2001	-2.54%	0.00
2002	-2.78%	0.00
2003	-3.04%	0.00
2004	-3.19%	0.01

Fonte: Cálculos dos autores.

Notas: No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização.

### 3.6. Crescimento da produtividade total de fatores

Nesta secção recorre-se aos parâmetros da função custo estimada para calcular uma medida de variação da produtividade total dos fatores (*TFPC*) que pode ser decomposta no efeito de variações na eficiência-custo (*EC*), do progresso tecnológico (*TP*), e da evolução dos retornos à escala (*RTS*) (para mais detalhes, ver Bauer (1990)):

$$TFPC = EC + TC + RTS \tag{14}$$

ou

$$\ln\left(\frac{TFP_{i,t}}{TFP_{i,t-1}}\right) = \ln\left(\frac{CE_{i,t}}{CE_{i,t-1}}\right) + \frac{1}{2}\left(-\frac{\partial \ln C_{i,t}}{\partial t} - \frac{\partial \ln C_{i,t-1}}{\partial t}\right) + \frac{1}{2}\sum_r \left[\left(\varepsilon_{r,i,t} \frac{1 - ESE_{i,t}}{ESE_{i,t}} + \varepsilon_{r,i,t-1} \frac{1 - ESE_{i,t-1}}{ESE_{i,t-1}}\right) \ln\left(\frac{y_{r,y,t}}{y_{r,y,t-1}}\right)\right] \tag{15}$$

Em que  $\varepsilon_r$  representa a elasticidade do custo em relação ao *output* *r* e cada um dos termos da

decomposição tem uma interpretação relevante. De facto, de acordo com a expressão apresentada acima, as variações da produtividade total dos fatores dos bancos podem ser determinadas pela aproximação à fronteira de custo (variações na eficiência-custo), por deslocamentos da própria fronteira ao longo do tempo (progresso tecnológico) e por movimentos ao longo da função custo (variações nas economias de escala). Este último efeito representa o impacto nos custos motivado diretamente por variações na atividade depois de permitir o ajustamento dos *inputs* e é positivo caso um banco com economias (deseconomias) de escala aumente (diminua) a sua atividade.

Note-se que a Equação (15) está especificada de forma semelhante à proposta em Bauer (1990), com as alterações necessárias para a adaptar a uma função custo que inclui o nível dos capitais próprios. Assim, recorre-se ao conceito de economias de escala em termos económicos e não à medida clássica de economias de escala apresentada na Equação (9).

Os resultados relativos à evolução da produtividade total dos fatores no setor bancário português durante o período em análise são apresentados no Quadro 8 e no Gráfico 5. Um resultado a salientar prende-se com o facto do crescimento da produtividade total de fatores ter sido determinado essencialmente pelo progresso tecnológico, que se intensificou ao longo do período da amostra. O efeito das economias de escala apresentou também um contributo positivo para o crescimento da produtividade, especialmente durante o período mais recente, uma vez que se verificou um aumento da atividade por parte de bancos cuja tecnologia exibia economias de escala<sup>14</sup>.

#### Quadro 8

CRESCIMENTO DA PRODUTIVIDADE TOTAL DE FATORES				
Ano	Evolução da Eficiência de Escala	Evolução da Eficiência Técnica	Progresso Tecnológico	Evolução da Produtividade Total de Fatores
	(1)	(2)	(3)	(4)
1993	0.39%	0.00%	-0.05%	0.34%
1994	0.36%	0.00%	0.21%	0.57%
1995	0.28%	0.00%	0.50%	0.78%
1996	0.17%	0.00%	0.82%	0.99%
1997	0.28%	0.00%	1.12%	1.41%
1998	1.00%	-0.01%	1.37%	2.37%
1999	0.73%	0.00%	1.60%	2.33%
2000	2.83%	0.00%	2.00%	4.83%
2001	1.33%	0.00%	2.38%	3.71%
2002	0.18%	0.00%	2.66%	2.83%
2003	0.44%	0.00%	2.91%	3.35%
2004	1.11%	0.00%	3.12%	4.23%

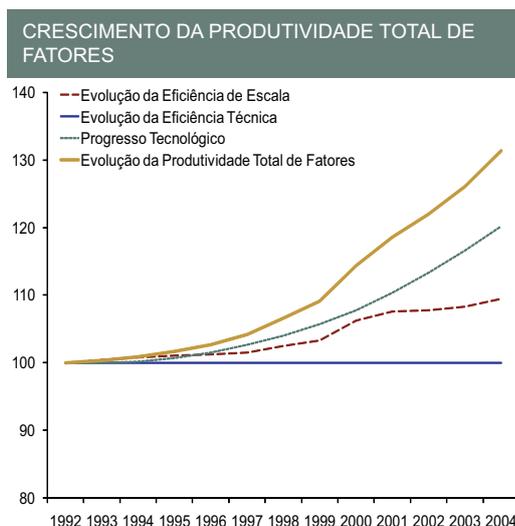
Fonte: Cálculos dos autores.

Notas: No cálculo das médias são utilizados como pesos os valores para os empréstimos concedidos ajustados para operações de titularização.

Durante o período em análise não se verificaram alterações significativas na eficiência-custo dos bancos portugueses. Conjugando os três efeitos, conclui-se que o aumento da produtividade foi reduzido durante o início da década de 90, mas intensificou-se consideravelmente ao longo do tempo, atingindo um valor de 4% em 2004.

(14) Este efeito foi particularmente marcado em 2000, ano em que se verificou um forte movimento de consolidação no setor.

Gráfico 5



Fonte: Cálculos dos autores.

#### 4. CONCLUSÕES

Neste artigo é desenvolvida uma análise da tecnologia de produção dos bancos portugueses ao longo do período compreendido entre 1992 e 2004, através da estimação de uma fronteira de custo *translog*. No modelo considerado, os bancos são empresas que produzem empréstimos e outros ativos que rendem juros, escolhendo a combinação dos níveis de mão-de-obra, capital físico e dívida que minimiza os custos, sujeito à detenção de um dado nível de capitais próprios.

Foram testadas algumas especificações alternativas para a distribuição da ineficiência estimada. Os resultados obtidos indicam que os bancos com risco de crédito mais elevado e com maiores reservas de liquidez tendem a ter níveis de ineficiência mais elevados, possivelmente refletindo o facto de estas variáveis funcionarem como indicadores da qualidade/sofisticação da gestão. Por outro lado, os bancos relativamente maiores e com menor alavancagem tendem a ser mais eficientes, o que poderá indicar, respetivamente, que os bancos de maior dimensão têm a capacidade de atrair equipas de gestão mais competentes e que uma capitalização mais robusta é uma forma eficaz de reduzir os problemas de agência. A análise mais detalhada cujos resultados são brevemente sintetizados em seguida foi desenvolvida com base nos resultados da estimação de um modelo mais simples no qual não foram incluídas quaisquer variáveis como determinantes da ineficiência.

As estimativas obtidas para os custos marginais dos bancos portugueses na produção de empréstimos e de outros ativos que rendem juros seguem em grande medida o comportamento das taxas de juro nominais ao longo do período em análise. Ainda assim, ter-se-á verificado um contributo não negligenciável para o comportamento do custo marginal total da redução do custo marginal não associado aos recursos financeiros. Em 2004, o último ano incluído neste exercício, a estimativa para esta medida cifrou-se em 1.2% na produção de empréstimos (o valor correspondente para os outros ativos que rendem juros é de 1.3%).

A estrutura de capital dos bancos foi considerada na análise através da inclusão do nível dos capitais próprios como um *input* fixo na função custo. Este procedimento permite o cálculo de estimativas para o custo sombra dos capitais próprios dos bancos, que deverá ser interpretado como um limite inferior à verdadeira disponibilidade dos bancos a pagar pelo capital. De facto, os valores encontrados para o custo sombra dos capitais próprios dos bancos são inferiores àqueles que seriam compatíveis com os valores geralmente considerados como razoáveis para o prémio de risco. Adicionalmente, esta estimativa segue de forma bastante próxima os desenvolvimentos nas taxas de juro de mercado.

Em média, os bancos portugueses operaram com um nível de ineficiência próximo de 9%, sugerindo que poderiam em teoria, mantendo o seu nível de atividade, incorrer apenas 91% dos custos observados. Verificou-se também ao longo período em análise uma aceleração do progresso tecnológico que se situou em 2.2% na média (ponderada) da amostra e em 3.2% em 2004. Tendo em conta a estrutura de capital dos bancos, foram identificadas economias de escala significativas, essencialmente no período mais recente. Adicionalmente, os resultados apontam para a existência de economias de gama na produção conjunta de empréstimos e de outros ativos que rendem juros.

No contexto da liberalização e privatização do sistema bancário e de crescente inovação financeira, a fronteira de custos que representa as melhores práticas deslocou-se no sentido descendente ao longo do tempo. A distância entre os custos observados dos bancos e a fronteira, por outro lado, não terá sofrido alterações significativas. O crescimento da atividade de bancos com rendimentos crescentes à escala implica um movimento ao longo da função custo que contribuiu também para o aumento da produtividade.

Conjugando estes resultados, foram obtidas estimativas para a produtividade total dos fatores na ordem de 2.8% ao ano em média, o que em termos acumulados resulta num aumento de produtividade de 31.4% entre 1992 e 2004. O aumento da produtividade foi mais marcado no período mais recente, atingindo um valor de 4.2% em 2004.

## REFERÊNCIAS

- Banco de Portugal (2009) “A economia portuguesa no contexto da integração económica, financeira e monetária”, Departamento de Estudos Económicos.
- Battese, G. E., e Coelli, T. J., (1992) “Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India”, *Journal of Productivity Analysis*, 3, 153-159.
- Battese, G. E, e Coelli, T. J., (1995) “A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data”, *Empirical Economics*, 20, 325-332.
- Bauer, P. W. (1990) “Decomposing TFP Growth in the Presence of Cost Inefficiency, Nonconstant Returns to Scale, and Technological Progress”, *Journal of Productivity Analysis*, 1, 287-299.
- Berger, A. N. e Mester, L. J. (1997) “Inside the black box: what explains differences in the efficiencies of financial institutions” *Journal of Banking & Finance*, 21, 895-947.
- Boucinha, M. e Ribeiro, N. (2009) “An assessment of competition in the Portuguese banking system in the 1991-2004 period”, Banco de Portugal, *Working Paper*, 1.
- Boucinha, M., Ribeiro, N. and Weyman-Jones, T. (2009) “An assessment of Portuguese banks’ costs and efficiency”, Banco de Portugal, *Working Paper* 22.
- Canhoto, A. e Dermine, J. (2003) “A note on banking efficiency in Portugal, New vs. Old banks”, *Journal of Banking and Finance*, Vol. 27, pp. 2087-2098.
- Freixas, X. e Rochet, J.-C. (1998) “Microeconomics of Banking”, MIT Press, Cambridge, London.
- Hughes, J. P. e Mester, L. J. (1993) “Accounting for the demand for financial capital and risk-taking in bank cost functions”, *Working Papers* 93-17, Federal Reserve Bank of Philadelphia.
- Hughes, J. P. (1999) “Incorporating Risk into the Analysis of Production,” *Atlantic Economic Journal*, 27:1, 1-23.
- Hughes, J. P., Mester, L. J. e Moon, C.-G. (2001) “Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production”, *Journal of Banking and Finance*, Elsevier, vol. 25 (12), pp. 2169-2208.
- Lima, F. (2008) “Product differentiation and the measurement of cost efficiency in banking: the case of Portuguese banks”, in Soares, J., Pina, J., e Catalão-Lopes, M. (eds), “New Developments in Financial Modelling”, Cambridge Scholars Publishing.
- Lima, F. e Pinho, P. S. (2008) “Financial disintermediation and the measurement of efficiency in banking: the case of Portuguese banks”, *International Journal of Banking, Accounting and Finance*, Vol. 1(2), pp. 133-148.
- Mendes, V. e Rebelo, J. (1999) “Productive efficiency, technological change and productivity in Portuguese banking”, *Applied Financial Economics*, Vol. 9, pp. 513-521.
- Mendes, V. e Rebelo, J. (2000) “The Effect of Bank M&As on Efficiency: The Portuguese Experience”, in Abreu, M. e Mendes, V. (eds), “What Financial System for Year 2000?”
- Pinho, P. S. (2001) “Using accounting data to measure efficiency in banking: an application to Portugal”, *Applied Financial Economics*, Vol. (11), pp. 527-538.

Ribeiro, N. (2007), "Debt growth: Factors, institutional issues and implications – The Portuguese case" in Enoch, C. e Ötker-Robe, I. (eds.), "Rapid credit growth in central and eastern Europe: Endless boom or early warning?" International Monetary Fund, 2007.

Sealey, C.W. e Lindley, J.T. (1977), "Inputs, Outputs and a Theory of Production and Cost at Depository Financial Institutions", *Journal of Finance*, Vol. 32(4), pp. 1251-1266.