UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

**A MENSURAÇÃO DO PRODUTO, EFICIÊNCIA E ECONOMIAS DE ESCALA DOS
BANCOS BRASILEIROS**

Thomas Fujiwara

Orientador: Prof. Dr. Márcio Issao Nakane

SÃO PAULO
2006

Profa. Dra. Suely Vilela
Reitor da Universidade de São Paulo

Profa. Dra. Maria Tereza Leme Fleury
Diretora da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Joaquim José Martins Guilhoto
Chefe do Departamento de Economia

Prof. Dr. Dantes Mendes Aldrighi
Coordenador do Programa de Pós-Graduação em Economia

THOMAS FUJIWARA

**A MENSURAÇÃO DO PRODUTO, EFICIÊNCIA E ECONOMIAS DE ESCALA DOS
BANCOS BRASILEIROS**

Dissertação apresentada ao Departamento de
Economia da Faculdade de Economia,
Administração e Contabilidade da Universidade
de São Paulo como requisito para a obtenção do
título de Mestre em Economia

Orientador: Prof. Dr. Márcio Issao Nakane

SÃO PAULO

2006

Dissertação defendida e aprovada no Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo – Programa de Pós-Graduação em Economia, pela seguinte banca examinadora:

“Tudo que precisa ser dito,
e tudo que deve ser entendido,
jamais caberiam em uma dedicatória.”

À Renata

AGRADECIMENTOS

Primeiramente, agradeço ao prof. Márcio I. Nakane pelo contínuo apoio, incentivo e orientação, que tanto contribuíram para a elaboração deste trabalho como para a formação de seu autor.

Pelas valiosas sugestões durante o exame de qualificação e por aceitarem o convite de participarem da banca de avaliação desta dissertação, agradeço a Ana Carla A. Costa e Danilo C. Igliori. Ademais, reitero meus agradecimentos à primeira que, juntamente com Benjamin M. Tabak, contribuíram com valiosos comentários ao projeto de pesquisa que originou esta dissertação.

Agradeço também aos professores da FEA-USP, pelos conhecimentos transmitidos, e aos funcionários da Coordenação de Pós-Graduação do IPE e da Secretaria de Pós-Graduação da FEA, pela constante presteza com que sempre me atenderam.

Diversos colegas do programa de mestrado do IPE-USP contribuíram de inúmeras maneiras para a realização deste trabalho. Para não cometer nenhuma injusta omissão, opto por não tentar listá-los.

Esta dissertação não teria sido possível sem a cooperação do Banco Central do Brasil, que gentilmente forneceu os dados utilizados.

Também merece agradecimento o apoio financeiro oferecido pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP a este projeto de pesquisa, assim como o apoio dado pela Fundação Instituto de Pesquisas Económicas – FIPE e Conselho Nacional para o Progresso Científico e Tecnológico – CNPQ durante os primeiros meses de meus estudos no IPE-USP.

Finalmente, agradeço a meu pai, que sempre apoiou incondicionalmente todas minhas escolhas, e à Renata, por tudo que fez e fará por mim.

RESUMO

Este trabalho aplica metodologia de Wang (2003a, 2003b) para definir uma nova medida do produto de bancos brasileiros. Acredita-se que tal medida seja superior às comumente utilizadas na literatura por se tratar de uma variável de fluxo, incorporar os depósitos bancários de maneira teoricamente embasada e levar em consideração a exposição ao risco. Esta nova variável de produção é utilizada na estimação de fronteiras estocásticas de produção e custo para o setor bancário brasileiro, visando a mensurar sua eficiência técnica e econômica, assim como a magnitude de suas economias de escala. As fronteiras estimadas apresentam especificação dada pela forma funcional flexível de Fourier e incorporam variáveis determinantes da eficiência. Os resultados apontam para uma acentuada ineficiência do setor bancário, assim como para a ocorrência de retornos crescentes de escala.

ABSTRACT

This work applies the Wang (2003a, 2003b) methodology to define Brazilian banks' output. It is believed that this new output measure is superior to the ones commonly used by the literature, since it treats output as a flow variable, provides a theoretical basis for defining the role of deposits and takes account of risk exposure in defining output. This new measure is applied to the estimation of stochastic production and cost frontiers for the Brazilian banking sector, aiming at measuring its technical and economic efficiency, and also the size of its scale economies. The frontiers follow a Fourier flexible functional form specification and incorporate efficiency determinants. The results point to the existence of high inefficiency in the banking industry, and also to the occurrence of increasing returns to scale.

SUMÁRIO

1 INTRODUÇÃO	3
2 MEDIDAS DE PRODUÇÃO BANCÁRIA	7
2.1 Metodologias de Mensuração da Produção Bancária	7
2.1.1 Abordagens baseadas em variáveis de estoque.....	7
2.1.2 Abordagens baseadas em variáveis de fluxo.....	12
2.2 Medidas de Produção Bancária na Literatura Brasileira.....	19
2.3 Cálculo do Produto Bancário	22
3 APLICAÇÃO PARA A ESTIMAÇÃO DA EFICIÊNCIA E ECONOMIAS DE ESCALA.....	27
3.1 Dados.....	27
3.2 Metodologia	30
3.3 Resultados: Fronteira de Produção.....	36
3.4 Resultados: Fronteira de Custo	39
3.5 Resultados: Determinantes da Ineficiência Técnica e Econômica.....	41
3.6 Resultados: Medidas de Economias de Escala e Ineficiência	45
4 CONCLUSÃO	49
REFERÊNCIAS.....	51
APÊNDICE 1 – CÁLCULO DAS VARIÁVEIS USADAS NO CAPÍTULO 3	55
APÊNDICE 2 – RESULTADOS AUXILIARES DA ESTIMAÇÃO DAS FRONTEIRAS DE PRODUÇÃO E CUSTO	59

LISTA DE QUADROS

Quadro 2.1: Representação Gráfica do Produto Implícito Bancário.....	18
Quadro 2.2: Metodologias para definição de produtos/insumos na literatura brasileira.....	20
Quadro 2.3: Mensuração do Produto Implícito Bancário	24
Quadro 3.1: Sumário da Amostra (em R\$ constantes de Dezembro/2004)	29
Quadro 3.2: Representação Gráfica da Eficiência Técnica e Alocativa.....	30
Quadro 3.3: Parâmetros da Fronteira de Produção	37
Quadro 3.4: Elasticidades da Fronteira de Produção	38
Quadro 3.5: Testes de Especificação da Fronteira de Produção	38
Quadro 3.6: Parâmetros da Fronteira de Custo	39
Quadro 3.7: Testes de Propriedades da Fronteira de Custo	40
Quadro 3.8: Testes de Especificação da Fronteira de Custo	41
Quadro 3.9: Parâmetros da Ineficiência Técnica (Fronteira de Produção)	43
Quadro 3.10: Parâmetros da Ineficiência Econômica (Fronteira de Custo).....	43
Quadro 3.11: Medidas de Eficiência e Economias de Escala	46

1 INTRODUÇÃO

O Brasil já conta com uma incipiente literatura envolvendo a estimação de funções de produção e custo para sua indústria bancária. Servem como exemplos aplicações à estimação da produtividade, eficiência, economias de escala e testes de concorrência.

Nestes estudos, assim como na literatura internacional, é prática comum definir o produto da atividade bancária como itens do ativo financeiro das instituições estudadas. Contudo, tal abordagem apresenta algumas deficiências, das quais é possível listar três como as mais relevantes.

Primeiramente, os itens do ativo de uma instituição financeira constituem um estoque, enquanto se espera, com base na teoria econômica, que o produto de uma determinada firma seja uma variável de fluxo.

Em segundo lugar, esta abordagem não clarifica o papel dos depósitos. Na literatura nacional, casos onde estes são considerados como produtos, insumos ou simplesmente relevados na análise são igualmente comuns. Em terceiro lugar, não é dada a devida atenção à exposição ao risco na atividade bancária.

Estas deficiências são em grande parte decorrentes da falta de embasamento teórico claro para o uso de ativos como medida da produção bancária. Visando superar isto, Wang (2003a, 2003b) apresenta modelo teórico que permite estabelecer os produtos e insumos da atividade bancária de maneira a superar as deficiências listadas acima.

Deste modo, este trabalho consiste em aplicar uma variação da metodologia de Wang (2003a, 2003b) para uma amostra de bancos brasileiros. Ademais, as séries estimadas são utilizadas para revisitar a questão, já estudada por Nakane (1999) e Silva e Jorge Neto (2002), da mensuração da eficiência e economias de escala na indústria bancária através da estimação de fronteiras estocásticas.

Além da introdução de nova medida de produto, este trabalho também adiciona dois novos fatores à literatura nacional. Enquanto Nakane (1999) e Silva e Jorge Neto (2002) estimam

fronteiras de custo com especificação *translog*, o presente trabalho trabalha com fronteiras de custo e produção que apresentam forma funcional de Fourier, que inclui a *translog* como um caso específico e permite uma aproximação mais flexível de fronteiras desconhecidas.

Neste sentido, cumpre ressaltar que, dentro do conhecimento do autor, este é o primeiro trabalho a utilizar a especificação de Fourier para analisar o setor bancário brasileiro, assim como também é o único a estimar tanto fronteiras estocásticas de custo como de produção.

Optou-se por fazer uma primeira aplicação desta nova medida da produção bancária na estimação da eficiência e economias de escala pois se acredita que esta seja especialmente sensível aos problemas do uso de ativos como medida de produção bancária.

Primeiramente, note-se que a eficiência técnica é definida como a capacidade de uma firma extrair a quantidade máxima de produto de um certo vetor de insumos, e que as economias de escala são mensuradas através da magnitude do impacto de variações da quantidade de insumos na quantidade produzida.

Assim, fica claro que tanto a mensuração da eficiência e de economias de escala são especialmente sensíveis à definição dos insumos e produtos, o que faz com que o uso de uma metodologia que defina o papel dos depósitos de maneira teoricamente embasada seja especialmente relevante.

Em segundo lugar, o uso de uma medida de produto que desconsidere a exposição ao risco na análise da eficiência pode acarretar viés significativo ao se “confundir eficiência com exposição ao risco”.

Um exemplo pode clarear este ponto. Assuma, por exemplo, que dois bancos (A e B) utilizem quantidades idênticas de insumos, sendo que o banco A empresta X reais a firmas que apresentam probabilidade nula de inadimplência, enquanto o banco B empresta $2X$ reais para firmas com probabilidade de inadimplência igual a 75%, sendo que ambos praticam a mesma taxa de juros.

Caso se considere que o produto dos bancos iguala seu ativo financeiro, o banco B produz o dobro do que o banco A . Além disto, por usarem a mesma quantidade de insumos, o banco B

é mais eficiente. Contudo, a receita esperada do banco A é duas vezes maior que a do banco B , de tal modo que este último é na realidade menos eficiente. Assim, o exemplo indica que o uso de uma medida de produto que considere a exposição ao risco de cada banco é crucial na análise de eficiência no setor bancário.

Este trabalho está dividido em quatro capítulos, incluindo esta introdução. Primeiramente, o Capítulo 2 apresenta as diferentes metodologias usadas para definir os produtos e insumos da atividade bancária e como estas foram aplicadas na literatura brasileira, assim como discute a aplicação da metodologia de Wang (2003a, 2003b) para o caso brasileiro.

O Capítulo 3, por sua vez, utiliza a metodologia discutida no capítulo anterior para estimar fronteiras estocásticas de produção e custo para um painel (com periodicidade mensal) de bancos brasileiros durante o ano de 2004. Um capítulo final conclui o trabalho.

2 MEDIDAS DE PRODUÇÃO BANCÁRIA

O objetivo deste capítulo é fazer uma rápida apresentação das metodologias de mensuração do produto bancário e apontar como estas foram usadas pela literatura brasileira. É também apresentada a metodologia de Wang (2003a, 2003b) e discutida uma maneira de aplicá-la ao caso brasileiro, através dos dados do Plano de Contas das Instituições Financeiras – COSIF, coletado pelo Banco Central do Brasil.

O capítulo divide-se em três partes. A primeira apresenta as metodologias para definir os produtos e insumos da atividade bancária. A segunda parte revisa brevemente a literatura envolvendo a estimação de funções de produção ou custo para a indústria bancária brasileira, com ênfase nas metodologias utilizadas para definir os produtos e insumos bancários.

Por fim, a Seção 2.3 trata da aplicação da metodologia de Wang (2003a, 2003b) para o caso brasileiro.

2.1 Metodologias de Mensuração da Produção Bancária

Inexiste consenso acerca da metodologia ideal para definir os produtos e insumos da atividade bancária, sendo que há grande variação nas abordagens usadas na literatura internacional e nacional. As abordagens existentes podem ser divididas em dois grupos distintos: as baseadas no valor contábil de ativos financeiros dos bancos (variáveis de estoque) e as que utilizam dados de fluxo.

Esta seção primeiramente apresenta as três abordagens mais comuns baseadas em variáveis de estoque. Em segundo lugar, são apresentadas duas metodologias que consideram o produto como fluxo, a baseada nas contas nacionais (mais especificamente no *System of National Accounts 1993* – SNA93 elaborado pela Organização das Nações Unidas – ONU) e a proposta de Wang (2003a, 2003b).

2.1.1 Abordagens baseadas em variáveis de estoque

Primeiramente, note-se que “*virtualmente todos estudos em organização industrial empírica usam o valor contábil de um tipo ou outro de ativo financeiro como medida de produto [bancário]*” (Wang, 2003a, tradução própria).

Contudo, são raros os estudos que buscam fundamentos teóricos para tanto. Uma das poucas tentativas está em Pesek (1970), que busca justificar o uso do valor contábil de itens dos balanços dos bancos como medida de produto. O argumento reside no fato de que os estoques financeiros do banco estão sempre “*deslizando no abismo da inexistência*” (Pesek, 1970, tradução própria), no sentido de que os depósitos são constantemente sacados e os empréstimos e títulos estão constantemente vencendo, o que faz com que o banco utilize recursos não-financeiros (trabalho e capital) para apenas manter a existência destes estoques.

Contudo, o fato dos bancos estarem continuamente “produzindo” a manutenção de seus estoques não implica necessariamente que o valor contábil destes seja uma medida do fluxo de produção relativa a esta manutenção. Assim, o argumento de Pesek (1970) supõe implicitamente que o fluxo de produção referente à manutenção dos estoques é proporcional ao valor destes.

Esta proporcionalidade é testada (e rejeitada) por Humphrey (1992), que verifica a correspondência entre medida de fluxo (número de transações de crédito e depósito registradas) e medidas de estoque para bancos norte-americanos. A análise de séries temporais agregadas para o período 1967-1987 mostra que a razão entre a medida de fluxo e a medida de estoque mais comum na literatura (valor das contas de depósito e crédito) apresenta grande queda (37%) no tempo, apesar das séries apresentarem certa correlação (0.82).

Ademais, a proporcionalidade é rejeitada mais fortemente na análise de Humphrey (1992) de uma *cross-section* de bancos. A correlação entre as séries é relativamente baixa (0.60) e grandes disparidades são observadas. Um exemplo é o fato de que, usando a medida de fluxo, o produto de um banco grande é em média 53 vezes maior que o de um pequeno, enquanto a utilização do valor de contas de depósito e crédito como produto indicaria um produto médio 219 vezes maior¹.

¹ Os bancos são classificados entre pequenos, médios e grandes de acordo com o tamanho de seus ativos.

Humphrey (1992) também aponta que o cômputo de medidas de produtividade total dos fatores dos bancos, seja por números-índice ou pela estimação de funções de produção, também são sensíveis à medida de produto utilizada, sendo que o próprio sinal da variação na produtividade estimada para o período estudado depende das medidas utilizadas.

Apesar do uso de contas do ativo e passivo como produto bancário ser comum na literatura, inexistente consenso no que tange o critério para definir se determinado item contábil constitui um produto ou insumo. Abaixo são discutidas as três abordagens mais usuais: a abordagem dos ativos (ou da intermediação), a abordagem do valor adicionado (ou abordagem da produção) e a abordagem do *user cost*.²

Abordagem dos ativos

A abordagem dos ativos é a mais simples e intuitiva das metodologias para definir a produção bancária. Ela considera que a atividade bancária está resumida na intermediação entre portadores de fundos emprestáveis e aqueles que recebem estes fundos.

Deste modo, esta abordagem considera os empréstimos e outros ativos financeiros como produto e os depósitos (e outros passivos financeiros) como insumos da atividade bancária. Por outro lado, alguns estudos desconsideram o papel dos depósitos e outros passivos enquanto insumo, tratando o banco como uma firma que utiliza combinações de trabalho e capital para produzir ativos.

Berger e Humphrey (1992) apontam que esta abordagem pode ser adequada em estudos acerca da lucratividade ou custos de empréstimos, baseados na estimação de formas reduzidas (não-estruturais).

A principal deficiência da abordagem reside no fato dos serviços prestados aos depositantes serem desconsiderados. Como Wykoff (1992) aponta, a noção de que depósitos têm apenas características de insumo implica que agentes econômicos racionais gastam recursos (tempo, locomoção e tarifas bancárias) para gratuitamente prover insumos valiosos aos bancos.

² Esta discussão é baseada em Berger e Humphrey (1992).

Ademais, Triplett (1992) coloca que a abordagem dos ativos sofre com problemas de arbitrariedade, no sentido de que a metodologia não apresenta mecanismos para que seja possível contestar os critérios usados para determinar se um grupo de ativos (passivos) deve ser considerado como produto (insumo) ou não.

Sobre isto, a aplicação deste conceito pode levar a contradições na prática, como demonstra um exemplo sugerido por Berger e Humphrey (1992). Suponha que o banco A empresta todos seus fundos (provenientes de depósitos) para o banco B, que por sua vez empresta este dinheiro às instituições não financeiras. Se empréstimos interbancários forem considerados como produto, então uma fusão entre os dois bancos diminuiria a produção conjunta destes, o que não faz sentido. Por outro lado, caso os empréstimos interbancários fossem desconsiderados, a produção do banco A seria zero, apesar de seus dispêndios com insumos não financeiros (trabalho e capital).

Abordagem do valor adicionado

A abordagem do valor adicionado propõe que as atividades responsáveis por grandes parcelas dos custos operacionais (insumos não financeiros) dos bancos sejam consideradas como produtos, enquanto as demais sejam consideradas como insumos ou produtos pouco importantes, a depender de outras características.

A justificativa da abordagem reside na hipótese de que as atividades responsáveis por partes consideráveis dos gastos em trabalho e capital são aquelas que apresentam maior valor adicionado.

Deste modo, a abordagem necessita de outra fonte de informação (afora os dados contábeis) que indique qual é a alocação dos custos operacionais dos bancos. Apesar de Berger e Humphrey (1992) utilizarem dados de natureza contábil que indicam a parcela dos custos operacionais pelo qual um determinado item contábil (depósitos a vista ou crédito imobiliário, por exemplo) é responsável,³ outras abordagens para determinar a alocação de custos operacionais podem ser usadas. Berger e Humphrey (1992) sugerem que os coeficientes de uma regressão linear do custo operacional total contra os valores dos itens financeiros do

³ Trata-se do programa de *Functional Cost Analysis* feito pelo Federal Reserve dos EUA para os bancos deste país.

ativo e passivos podem ser interpretados como a porcentagem do custo operacional pelo qual cada item é responsável.

A aplicação da metodologia por Berger e Humphrey (1992) ao caso norte-americano indica duas categorias de depósitos como sendo responsáveis por 48% do valor adicionado, o que contraria fortemente a abordagem dos ativos, que computaria estas como sendo insumos à atividade bancária.

Contudo, faltam fundamentos teóricos sólidos para sustentar que a alocação dos custos operacionais seja uma *proxy* adequada para o valor adicionado e que este realmente identifique a distinção entre produtos e insumos, uma vez que é plausível que quantidades significativas de capital e trabalho sejam alocadas na obtenção e tratamento de insumos financeiros (Triplett, 1992).

Abordagem do user cost

Wang (2003a) aponta que, das três abordagens discutidas nesta subseção, a do *user cost* é a única que apresenta fundamentação em teoria microeconômica. O conceito do *user cost* é definido como a diferença entre o custo de oportunidade e a receita (custo) de um ativo (passivo).

Deste modo, o trabalho pioneiro de Hancock (1985) postulou que categorias contábeis possuindo *user cost* positivo devem ser consideradas como insumos, enquanto as que apresentam *user cost* negativo seriam produtos.

A aplicação da metodologia de *user cost* apresenta sérias dificuldades de ordem prática. Primeiramente, a mensuração de receitas implícitas torna-se crucial não apenas na determinação do valor das variáveis em questão, mas também na definição do que virá a ser considerado como produto ou insumo.

Em segundo lugar, a definição do custo de oportunidade pode ser extremamente complexa, uma vez que variáveis específicas a cada firma bancária, como risco, liquidez e maturidade de seus ativos e passivos, devem ser consideradas adequadamente. Fixler e Zieschang (1992) mostram que inexiste uma metodologia consensual de definição dos custos de oportunidades,

sendo este problema agravado pelo fato dos métodos existentes levarem a resultantes diversos e muitas vezes conflitantes.

2.1.2 Abordagens baseadas em variáveis de fluxo

Como já apontado anteriormente, a construção das medidas de produto e insumo a partir de variáveis de fluxo é incomum na literatura em organização industrial empírica relacionada à indústria bancária. Deste modo, serão tratadas nesta seção a abordagem utilizada pelos sistemas de contas nacionais e a metodologia proposta por Wang (2003a, 2003b).

Abordagem das contas nacionais

No SNA93⁴, os bancos são vistos como firmas prestadoras de serviços de intermediação financeira. Dentro da lógica de mensuração das contas nacionais, o produto de uma firma de serviços pode ser medido pelas taxas e comissões cobradas. Contudo, é sabido que os bancos podem cobrar implicitamente por seus serviços nas taxas de juros cobradas (pagas) em seus empréstimos (depósitos).

Deste modo, United Nations (1993) coloca que a mensuração da parcela dos serviços bancários cobrada explicitamente “*não constitui problema conceitual ou prático*” (United Nations, 1993, tradução própria), residindo na parcela cobrada implicitamente a diferença entre a mensuração do produto bancário em relação ao produto de demais firmas prestadoras de serviços.

O valor total dos produtos de todas instituições financeiras⁵ (não apenas bancos comerciais) cobrado explicitamente é definido como sendo a diferença entre o total dos rendimentos de propriedade a receber e os juros totais a pagar. Esta medida é denominada de SIFIM – Serviços de Intermediação Financeira Indiretamente Medidos.

⁴ Compete à ONU estabelecer os padrões dos sistemas de contas nacionais adotados por diversos países (como o Brasil). Em 1993 foi feita a última grande reformulação destas normas, estabelecidas no SNA93 (*System of National Accounts 1993*) e detalhadas em United Nations (1993).

⁵ Com exceção de companhias de seguros.

O SNA93 aponta que o SIFIM (medido pela ótica do produto) deve ser alocado (pela ótica do dispêndio) entre o consumo intermediário das firmas, consumo final das famílias ou exportação⁶. Para tanto, é necessário que o valor adicionado dos bancos em suas diferentes operações de crédito e depósitos seja identificado.

Para identificar este valor adicionado, o SNA93 aponta que uma taxa de juros de referência deve ser usada. Assim sendo, o produto proveniente de uma operação de crédito é a diferença entre a taxa de juros efetivamente cobrados e aquela que seria cobrada se a taxa de referência fosse cobrada. Por outro lado, o produto proveniente de um depósito é a diferença entre os juros que seriam pagos se a taxa de referência fosse praticada e os juros efetivamente recebidos.

O SNA93 aponta que a taxa de referência deve ser o “*custo puro de tomar fundos emprestados*” (United Nations, 1993, tradução própria). Este custo seria dado por uma taxa de juros onde todo o prêmio de risco diversificável é eliminado e líquida de todo custo de intermediação. United Nations (1993) sugere que a taxa de empréstimos interbancários ou a taxa básica estabelecida pelo banco central sejam usadas.

Wang (2003a) critica a metodologia do SNA93 apontando que ela superestima o produto implícito dos empréstimos dos bancos. Este ponto será discutido adiante, quando sua proposta de medida do produto bancário for apresentada. Sobre isto, a autora também aponta que a abordagem do SNA93 levaria a um produto negativo (relativo à atividade de empréstimos) quando a taxa de juros dos empréstimos estiver abaixo da taxa de referência⁷.

Abordagem de Wang (2003a, 2003b)

Com base em modelo dinâmico onde os bancos maximizam o valor presente de seu fluxo futuro de lucros, Wang (2003a) estabelece uma nova metodologia para determinar o produto de uma firma bancária.

⁶ Contudo, devido às dificuldades na realização desta alocação com metodologia de significado econômico razoável, o SNA93 permite que países aloquem o SIFIM em uma única indústria fictícia com valor adicionado igual a zero. Esta alternativa é adotada no Brasil, que aloca todo SIFIM no Dummy Financeiro (IBGE, 2004).

⁷ Wang (2003a) argumenta que isto ocorre em momentos de recessão, quando as taxas de longo prazo caem abaixo da de curto prazo. Cumpre também notar que, neste caso, os bancos de fato incorrem perdas financeiras em sua atividade de crédito, porém prejuízo não significa de forma alguma produto negativo.

A abordagem consiste em dividir os bancos em três divisões hipotéticas, cada uma responsável por uma de três atividades bancárias distintas: serviços ofertados a tomadores de crédito, intermediação de fundos emprestáveis e serviços para depositantes.

A primeira divisão engloba os serviços relacionados à solução (ainda que parcial, em boa parte dos casos) de problemas de informação assimétrica associados ao mercado de crédito (monitoramento e *screening* dos tomadores de crédito). Deste modo, é possível ver esta área como uma firma que faz a certificação da qualidade dos tomadores de crédito.

A segunda divisão é responsável pela intermediação financeira “pura”, líquida de serviços como os relacionados à solução de problemas de informação assimétrica. As transferências de fundos emprestáveis entre poupadores e tomadores de crédito, assim como as transações de títulos, ações e derivativos, uma vez desconsiderados eventuais serviços prestados juntamente com estas operações, pertencem a este grupo.

Finalmente, a terceira divisão é responsável pela oferta de serviços para depositantes. Dentre eles, podemos citar serviços de consultoria sobre investimentos financeiros, a provisão de local seguro para os depositantes guardarem suas poupanças e de meios de pagamento (cartões de crédito/débito, transferências entre contas, pagamentos de contas, cheques, caixas eletrônicos, entre outros), assim como a realização da contabilidade (*bookkeeping*) do dinheiro depositado.

A primeira e terceira divisões utilizam recursos não financeiros (trabalho, capital e outros insumos) para produzir valor adicionado. Por outro lado, a segunda divisão não cria valor adicionado além daquele que seria esperado pelos investidores ao investir diretamente nos projetos intermediados. Um exemplo pode clarificar a questão: quando um depósito em um fundo (de títulos públicos, por exemplo) administrado por um banco é feito, este provê ao depositante serviços de consultoria financeira, *bookkeeping* dos recursos depositados e seus rendimentos, além de estar guardando a poupança do depositante (sob a forma de títulos). Estes serviços estão englobados na terceira divisão. Contudo, a operação de compra dos títulos públicos com os fundos do depositante não gera valor adicionado. Sobre isto, os rendimentos provenientes destes títulos constituem valor adicionado do depositante, e não do banco.

Deste modo, Wang (2003a) entende o produto da segunda divisão como um insumo intermediário, semelhante ao valor aos produtos fretados por transportadoras ou revendidos por supermercados. Assim, o produto da segunda divisão deve ser incluído em medidas do produto bruto, mas não em medidas de valor adicionado.

Wang (2003a) também chama a atenção para o fato de que o produto bancário relacionado à realização do *matching* entre ofertantes e demandantes de fundos emprestáveis também está contemplado nas divisões acima⁸. O valor deste serviço ao ofertante de fundos emprestáveis (depositante) está compreendido na terceira divisão, enquanto o valor deste serviço para o demandante (tomador de crédito) está compreendido na primeira divisão.

Cumprе notar que os bancos provêm serviços à medida que permitem que transações de ativos financeiros (a compra de um título público, por exemplo) sejam feitas sem que seu cliente incorra os custos de transação que ocorreriam caso a transação fosse feita fora do sistema bancário (o custo de fazer a compra diretamente no Tesouro Nacional). Apesar de não estar explicitamente apontado em Wang (2003a), sua metodologia contempla estes serviços, seja na primeira (quando este é fornecido ao tomador de crédito) ou na terceira (quando fornecido a um depositante).

Para evidenciar o grau de abrangência das três divisões acima, uma comparação com o trabalho de Levine (1997), que enumera as funções do sistema financeiro, pode ser útil. O trabalho aponta que as funções do sistema financeiro são:

- a) facilitar o comércio, diversificação e *pooling* do risco;
- b) facilitar o comércio de bens e serviços;
- c) adquirir informações sobre projetos e alocar recursos;
- d) monitorar gerentes e exercer controle corporativo; e
- e) mobilizar poupança.

⁸ Na ausência de um sistema financeiro, os poupadores teriam de incorrer em custos de transação para “encontrar” outros agentes com projetos necessitando financiamento. Ademais, projetos maiores podem requerer fundos emprestáveis de vários poupadores, sendo que o sistema financeiro (bancos) permite que um “*pooling* de poupadores” seja feito. Deste modo, o serviço de *matching* entre poupadores e despoupadores, ou a criação de um mercado de fundos emprestáveis, é um dos serviços bancários.

No que tange aos bancos, as funções (c) e (d) estão claramente compreendidas na primeira divisão. A função (b) está compreendida na terceira divisão, uma vez que os serviços pela provisão de caixas eletrônicos, cheques e meios de pagamentos eletrônicos estão incluídos nesta divisão. Pela discussão acerca da realização do serviço bancário de criação de um mercado de fundos emprestáveis, a função (e) está incorporada na primeira e terceira divisões.

Por fim, cumpre notar que a função (a), referente ao comércio, diversificação e *pooling* do risco, está incluída na segunda divisão. Isto ocorre pois o papel dos bancos nesta função está apenas em fazer a intermediação entre agentes que desejam adquirir ou vender ativos financeiros (títulos, ações, debêntures) com o intuito de alterar sua exposição ao risco⁹.

O fato da metodologia de Wang (2003a) ser capaz de abranger as cinco funções propostas por Levine (1997) evidencia suas vantagens em relação às outras metodologias de definição de produto bancário. A abordagem dos ativos (intermediação), por exemplo, claramente desconsidera parte do papel dos bancos nas funções de mobilização de poupança (o serviço de *matching* provido ao poupador) e facilitação do comércio (a criação de meios de pagamento).

Neste sentido, o uso da abordagem dos ativos ou de *user cost* provavelmente relevará algumas das funções acima, a depender da identificação dos itens contábeis definidos como insumos que os dados utilizados apresentarem.

Wang (2003b) descreve como aplicar a metodologia predominantemente teórica de Wang (2003a) na prática. Primeiramente, é preciso notar que, assim como é feito na abordagem das contas nacionais, o valor de um produto é definido como sua receita total. Deste modo, caso não houvesse receitas implícitas nas taxas de juros praticadas pelos bancos, o produto bancário seria facilmente medido pela receita com tarifas.

Assim, a explicação abaixo foca na mensuração das receitas implícitas da atividade bancária. Cumpre também notar que a mensuração dos valores nominais é discutida abstendo-se da questão acerca do deflator adequado para os diferentes produtos bancários.

⁹ Wang (2003a) também aponta que o serviço dos bancos em administrar a sua própria exposição ao risco deve ser entendido como um insumo intermediário produzido dentro da própria firma, e que portanto não afeta a mensuração do produto bancário.

O produto da primeira divisão é então definido como o total de juros recebidos subtraído do retorno esperado pela carteira de empréstimos do banco. Isto é o produto da primeira divisão (Y^1) é definido como:

$$Y^1 = [i^B - E(r^B)]B = \Delta i^B B \quad (1)$$

Onde i^B e $E(r^B)$ são, respectivamente, a taxa de juros efetivamente praticada e a taxa de retorno esperado, dado o risco sistemático, da carteira de empréstimos cujo saldo total é B . A lógica desta definição é que Δi^B representa a “margem” que o banco cobra sobre a taxa de juros esperada de seus empréstimos, $E(r^B)$, pelos serviços prestados aos tomadores de crédito. Contudo, a mensuração desta na prática pode ser complexa, dado que é difícil observar $E(r^B)$.

Wang (2003a) aponta que a maneira ideal de mensurar $E(r^B)$ é através da aplicação de modelos de precificação de ativos à carteira dos bancos. Contudo, quando esta opção não é possível, as taxas de juros de títulos comparáveis podem servir como *proxies* para $E(r^B)$.¹⁰

De maneira análoga, o produto implícito da terceira divisão é igual à diferença entre o pagamento de juros aos depositantes e o montante que os depositantes receberiam caso seus fundos rendessem a taxa de mercado. Isto é, o produto da terceira divisão (Y^3) é dado por:

$$Y^3 = \sum_m [r^F - r^m] D^m = \sum_m \Delta r^m D^m \quad (2)$$

Onde r^m e r^F são, respectivamente, a taxa de juros paga e a taxa de juros “de mercado” (no caso onde o depósito é garantido, esta é a taxa livre de risco) do depósito m , cujo valor é D^m . Assim como no caso do produto da primeira divisão, o termo r^F não é observado, e alguma taxa de juros deve ser usada como *proxy*. Wang (2003b) usa a taxa de juros do *Treasury Bill* de 90 dias para tanto.

¹⁰ Para estudar o caso dos EUA, Wang (2003b) usa como *proxy* o retorno do *commercial paper* de 90 dias para aproximar o retorno esperado de “*C&I, agricultural and consumer loans and lease financing receivables*”. A *federal funds rate* é usada para o caso de empréstimos para financiar operações com títulos e a taxa de hipoteca da *Federal Housing Administration* é aplicada ao crédito imobiliário.

Por fim, o produto da segunda divisão é definido como o fluxo de fundos emprestáveis transferidos em um certo período. Wang (2003b) aproxima este fluxo com uma variável de estoque: o valor total dos empréstimos feitos em determinado momento do tempo.

O Quadro 2.1 apresenta uma representação gráfica baseada em Wang (2003a, 2003b) que permite sumariar sua proposta. É representado um caso simplificado onde se supõe que o total emprestado iguala o total depositado e que a taxa que o banco remunera seus depósitos é inferior à taxa livre de risco.

Quadro 2.1: Representação Gráfica do Produto Implícito Bancário

Fonte: baseado em Wang (2003a)

Na figura apresentada, o produto da primeira divisão é representado pela área I e o produto da terceira divisão é representado pela área IV, estas duas áreas constituem valor agregado pelo banco, à medida que elas são receitas dos serviços prestados aos tomadores de crédito e depositantes. Por sua vez, a área II representa o prêmio de risco da carteira dos bancos, enquanto as áreas V e III representam, respectivamente, a remuneração dos depósitos e a contribuição para o seu seguro (no caso brasileiro, para o Fundo Garantidor de Crédito – FGC). Note-se que estas três áreas (II, III e V) não constituem valor adicionado dos bancos.

O Quadro 2.1 também permite demonstrar as diferenças entre a medida proposta e as existentes. A abordagem dos ativos consideraria como produto o comprimento horizontal do retângulo, que obviamente não guarda nenhuma proporção com as áreas II e IV. A medida das contas nacionais, por sua vez, usaria a taxa de juros “de mercado” para determinar o produto. Assim, a receita implícita do banco com empréstimos e depósitos seria, respectivamente, igual à soma das áreas I e II e a área III. Deste modo, a metodologia das contas nacionais superestima o produto mensurado segundo Wang (2003a, 2003b) no valor igual à área II.

2.2 Medidas de Produção Bancária na Literatura Brasileira

Apesar de recente, a literatura empírica envolvendo o setor bancário brasileiro é relativamente desenvolvida, principalmente quando comparada com aquelas envolvendo outras indústrias no país.

Esta seção revisa as metodologias de definição dos insumos e produtos da atividade bancária utilizadas pela parcela desta literatura onde funções (ou fronteiras) de produção ou custo são estudadas. Assim, estudos baseados na estimação de “formas reduzidas”, usualmente focados em áreas de atuação específicas dos bancos, assim como trabalhos estimando a demanda por produtos bancários, são excluídos da análise.

Por outro lado, testes de concorrência dentro do paradigma da *New Empirical Industrial Organization* (NEIO) foram incluídos nesta resenha, dado que estes estudos usualmente envolvem a estimação de uma relação de oferta com uma forma funcional específica para o custo marginal e exigem a definição de produto para a atividade bancária.

A base de dados usada por todos estudos revisados é o Plano de Contas das Instituições Financeiras – COSIF, elaborado pelo Banco Central do Brasil. Sobre isto, o Quadro 2.2 aponta que a abordagem dos ativos é a mais usada, apesar da definição de que contas específicas do ativo (passivo) são usadas para determinar os produtos (insumos) dos bancos apresente diferença entre os trabalhos.

Quadro 2.2: Metodologias para definição de produtos/insumos na literatura brasileira

<i>Estudos sobre Eficiência</i>			
Trabalho	Metodologia para Mensurar Eficiência	Definição de Produto	Papel dos Depósitos
Nakane (1999)	Fronteira Estocástica	Ativos financeiros	Insumo quasi-fixo
Campos (2002)	<i>Data Envelopment Analysis</i> (DEA) ¹¹	Ativos financeiros e depósitos à vista	Produto
Silva e Jorge Neto (2002)	Fronteira Estocástica	Abord. dos ativos	Insumo
Sousa, Staub e Tabak (2003)	<i>Data Envelopment Analysis</i> (DEA)	Ativos financeiros e depósitos à vista	Produto
Krause e Tabak (2004)	<i>Data Envelopment Analysis</i> (DEA)	IVA	Insumo
Chabalgoity, Marinho, Benegas e Jorge Neto (2005)	<i>Data Envelopment Analysis</i> (DEA)	Ativos financeiros	Insumo
<i>Outros Estudos</i>			
	Objetivo do Estudo	Definição de Produto	Papel dos Depósitos
Nakane (2002)	Teste de concorrência	Ativos financeiros	Nenhum
Belaish (2003)	Teste de concorrência	Variável de fluxo	Insumo
Costa e Nakane (2004)	Decomposição de <i>spread</i>	Ativos financeiros	Insumo
Nakane e Weintraub (2005)	Produtividade	Ativos financeiros	Nenhum
Jorge Neto, Araújo e Ponce (2005)	Teste de concorrência	Variável de fluxo	Insumo

Dentre os nove estudos que usam a abordagem dos ativos, apenas quatro aplicam sua versão “padrão”, considerando depósitos como insumos. Nakane (1999), Silva e Jorge Neto (2002), Costa e Nakane (2004) e Chabalgoity *et alli* (2005) consideram depósitos como insumos da atividade bancária. Contudo, os itens contábeis e tipos de depósitos que são considerados variam entre estudos¹². Note-se que Costa e Nakane (2004) e Chabalgoity *et alli* (2005) consideram as receitas de serviços como serviços, de tal modo que o produto bancário definido nestes estudos é um misto de variáveis de fluxo e estoque.

¹¹ Trata-se de método não-paramétrico para construir uma fronteira eficiente. Coelli, Battese e Rao (1998) apresentam uma introdução a esta metodologia.

¹² Sobre isto, vale notar que enquanto Nakane (1999) considera os depósitos como insumos quasi-fixos, os demais estudos os tratam como insumos variáveis.

Apesar de usarem uma variável de fluxo como produto (receita total), Belaisch (2003) e Jorge Neto, Araújo e Ponce (2005) também consideram fundos captados (despesas de captação) como insumos para realizar testes de concorrência.

Outros trabalhos dentro da abordagem dos ativos não consideram os fundos captados pelos bancos como insumos. Para mensurar a produtividade dos bancos, Nakane e Weintraub (2005) estimam uma função de produção onde as variáveis dependentes são apenas capital, trabalho e insumos intermediários (telefonias, eletricidade, entre outros). Nakane (2002) também não considera depósitos como insumo para testar a concorrência no setor bancário.

Elementos da abordagem do valor adicionado são incorporados por Campos (2002) e Sousa, Staub e Tabak (2003), que consideram depósitos à vista como produtos. Apesar de não aplicar a metodologia do valor adicionado propriamente, ambos estudos justificam a inclusão dos depósitos à vista à luz da literatura internacional baseada nesta abordagem¹³.

Destoa das demais metodologias a abordagem de Krause e Tabak (2004), que estima a fronteira DEA onde o produto é modelado pelo “valor intrínseco da empresa”, medida baseada no patrimônio e lucro líquido dos bancos que busca medir o valor presente de seus lucros futuros. Deste modo, Krause e Tabak (2004) calculam uma fronteira de lucro e não de produção, notando que o uso de uma técnica que não exija pressupostos acerca da forma funcional da fronteira (DEA) permite que esta ambigüidade não afete os resultados do trabalho.

Por outro lado, Krause e Tabak (2004) estabelecem como insumos trabalho, capital e recursos disponíveis para empréstimos. A inclusão deste último item como insumo faz com que a abordagem dos ativos seja aquela que, das descritas na Seção 2.1, mais se aproxima à utilizada.

Assim, conclui-se que existe grande variação nas metodologias para definir os produtos e insumos da atividade bancária, sendo que o item mais controverso aparenta ser o papel dos depósitos.

¹³ Como já apontado na Seção 2 deste capítulo, os depósitos à vista são um dos principais responsáveis pelo valor adicionado dos bancos norte-americanos (Berger e Humphrey, 1992).

2.3 Cálculo do Produto Bancário

Esta seção apresenta o procedimento realizado para calcular, com base na metodologia de Wang (2003a, 2003b) e nos dados do Plano Contábil das Instituições Financeiras – COSIF, coletados pelo Banco Central do Brasil, o produto de bancos múltiplos e comerciais em operação no Brasil.

O COSIF disponibiliza dados mensais sobre os saldos de operações de créditos e depósitos realizados nas instituições financeiras em operação no país, assim como traz informações sobre as receitas e despesas relacionadas a estas operações e aos serviços cujas taxas foram explicitamente cobradas pela instituição.

Deste modo, os dados permitem facilmente a identificar, através do valor da conta “*Rendas de Prestação de Serviços*” a parcela “explícita” do produto bancário, qual seja, aquela que é coletada através de taxas específicas para serviços aos tomadores de crédito e depositantes.

Ademais, os dados também permitem o cálculo, através da Equação (2), do produto implícito obtido junto aos depositantes. Este é definido pela multiplicação do saldo de depósitos e pelo diferencial entre a taxa de juros livre de risco e a taxa praticada pelo banco em seus depósitos. Assim, o valor da Equação (2) pode ser calculado considerando a taxa Selic como a livre de risco e aproximando a taxa com a qual o banco remunera seus depósitos através da divisão das despesas de captação pelo saldo de depósitos.¹⁴

Por outro lado, os dados do COSIF não permitem a observação direta de todos elementos da Equação (1), que define o valor do produto implícito obtido junto aos tomadores de crédito. Conseqüentemente, algumas aproximações baseadas no esquema apresentado no Quadro 2.1 foram realizadas.

Conforme apontado na Equação (1) – reapresentada abaixo – o produto implícito da primeira divisão hipotética de Wang (2003a) pode ser obtido pela multiplicação do saldo de

¹⁴ Considerou-se como “depósitos” os depósitos à vista, depósitos a prazo e cadernetas de poupança (ver Apêndice 1).

empréstimos do banco (B) pelo diferencial entre sua taxa de juros praticada (i^B) e a taxa de retorno esperado ($E(r^B)$) de sua carteira de empréstimos.

$$Y^l = [i^B - E(r^B)]B = \Delta i^B B \quad (1)$$

Wang (2003a, 2003b) aponta que uma das deficiências de sua metodologia é que $E(r^B)$ não é diretamente observável e apresenta uma grande exigência de dados para ser estimado de maneira “ideal” (através da aplicação de um modelo de apreçamento dos ativos na carteira de empréstimos dos bancos). Ademais, a solução encontrada por Wang (2003b) para este problema – o uso de taxas de juros de mercados de debêntures como *proxy* para $E(r^B)$ – seria pouco adequada ao caso do Brasil, onde os títulos de dívida privados têm um reduzido papel nos mercados de crédito.

Apesar de $E(r^B)$ não ser diretamente observável, o Quadro 2.1 aponta que este iguala a soma da taxa de juros livre de risco, a taxa de remuneração do seguro sobre depósitos e o prêmio de risco (de inadimplência). Destes, apenas o prêmio de risco não é observável, dado que se pode definir a taxa Selic como a livre de risco e a remuneração do seguro sobre depósitos no Brasil pode ser facilmente calculada, dado que o Fundo Garantidor de Crédito – FGC aplica uma alíquota uniforme de 0.025% a.m. nos depósitos cobertos pela garantia.

Assim, o principal obstáculo para o cálculo de $E(r^B)$ é a mensuração do prêmio de risco da carteira de empréstimos dos bancos. Contudo, uma possível solução seria utilizar a abordagem adotada por Banco Central do Brasil (2006) para a realização da decomposição do *spread* das operações de crédito bancário. Tal metodologia utiliza o saldo dos empréstimos atrasados como *proxy* da inadimplência.

A metodologia consiste em supor que 40% do valor dos empréstimos com mais de 90 dias de atraso seja uma boa medida da quantidade de empréstimos que serão inadimplentes na carteira de crédito de determinado banco.

Assim, o prêmio de risco r^{RP} pode ser aproximado pelo adicional de juros sobre a taxa livre de risco que faz com que sua carteira de crédito (que sofrerá determinada inadimplência) tenha retorno esperado igual à taxa de juros livre de risco. Ou seja:

$$(B - I)(r^F + r^{RP}) = Br^F$$

Onde r^F é a taxa de juros livre de risco e I é o saldo de empréstimos inadimplentes que, no caso, iguala 40% dos empréstimos com mais de 90 dias de atraso.

Rearranjando a equação acima, temos a equação que para o cálculo do prêmio de risco:

$$r^{RP} = \frac{Br^F}{(B - I)} - r^F \quad (3)$$

O Quadro 2.3 apresenta a fórmula para o cálculo do produto implícito dos bancos para as duas divisões hipotéticas de Wang (2003a).

Quadro 2.3: Mensuração do Produto Implícito Bancário

Divisão 1 – Crédito	Divisão 3 – Depósitos
Rendas de Operações de Crédito	(Taxa Selic)*(Saldo de Depósitos)
- (Taxa Selic)*(Saldo de Operações de Crédito)	
- (Alíquota FGC)*(Saldo de Operações de Crédito)	- Despesas de Captação
- (Prêmio de Risco)*(Saldo de Operações de Crédito)	
= Produto Implícito da Divisão 1	= Produto Implícito da Divisão 3

A soma destes dois produtos implícitos com o explícito iguala o valor adicionado do banco em questão. O Apêndice 1 apresenta os itens contábeis utilizados no cálculo em maior detalhe.

Por fim, é necessário realizar três ressalvas no que tange a aplicação prática da metodologia. Primeiramente, o esquema apresentado no Quadro 2.1 apresenta o caso onde a taxa praticada nos depósitos é inferior à taxa livre de risco. Contudo, isto pode não ocorrer na prática. Nestes casos, o produto da divisão 3 será negativo, o que pode ser interpretado como uma realização de subsídio cruzado entre o produto implícito e o produto explícito para depositantes.

Neste sentido, é importante notar que, mesmo nos casos onde o produto da divisão 3 é negativo, a taxa de retorno esperada ($E(r^B)$) dos empréstimos dos bancos deve ser aproximada pela soma da taxa Selic, a remuneração do FGC e a taxa de inadimplência. Caso a taxa Selic

fosse substituída pela taxa de remuneração de depósitos, haveria uma dupla contagem do produto (negativo) da divisão 3, dado que seu valor (em módulo) estaria sendo subtraído do produto da divisão 1.

Em segundo lugar, também é possível que a taxa de inadimplência seja negativa, sendo esta uma das principais deficiências práticas da metodologia adotada pelo Banco Central do Brasil. Contudo, como não há sentido em um prêmio de risco negativo, as taxas de inadimplência negativas foram substituídas por uma taxa igual a zero durante a realização dos cálculos.

Em terceiro lugar, note-se que, idealmente, os saldos de operações de crédito e depósitos a serem usados no cálculo devem ser os saldos médios do período analisado. Isto é, ao calcular o produto de um determinado semestre, deve-se usar o saldo médio deste semestre, e não o saldo observado no final do período. Contudo, o COSIF apresenta apenas os saldos em final de período.

3 APLICAÇÃO PARA A ESTIMAÇÃO DA EFICIÊNCIA E ECONOMIAS DE ESCALA

Este capítulo utiliza a metodologia de cálculo do produto bancário descrita na Seção 2.3 na estimação de fronteiras estocásticas de produção e custo para uma amostra de 77 bancos múltiplos e comerciais, visando mensurar a eficiência e economias de escala na indústria bancária brasileira.

Conforme apontado no Capítulo 1 deste trabalho, acredita-se que uma medida que considere a exposição ao risco e proponha um embasamento teórico claro para o papel dos depósitos na produção bancária seja de especial interesse no que tange a mensuração da eficiência e economias de escala.

Ademais, a literatura brasileira já conta com dois trabalhos que estimam fronteiras estocásticas de custo para a indústria bancária brasileira: Nakane (1999) e Silva e Jorge Neto (2002). Assim, a metodologia e os resultados deste trabalho serão contrastados com os destes estudos.

Este capítulo divide-se em seis partes. A Seção 3.1 apresenta os dados utilizados, enquanto a Seção 3.2 trata da metodologia de estimação. As Seções 3.3 e 3.4 discutem, respectivamente, os resultados da estimação das fronteiras de produção e de custo, deixando a análise dos determinantes da eficiência para a Seção 3.5 e das medidas de eficiência e economias de escala para a Seção 3.6.

3.1 Dados

Esta seção apresenta os resultados da aplicação da metodologia descrita na Seção 3.1 para calcular o produto mensal durante o ano de 2004 para uma amostra de 77 instituições financeiras caracterizadas pelo Banco Central do Brasil como bancos múltiplos ou comerciais.¹⁵

¹⁵ Incluindo o Banco do Brasil e a Caixa Econômica Federal.

Assim, a amostra constitui um painel balanceado de 12 períodos. Note-se também que os indivíduos da amostra são instituições financeiras, que podem fazer parte de um conglomerado financeiro. Ademais, há casos de duas (ou mais) instituições da amostra que pertencem ao mesmo conglomerado.

Apesar de 164 bancos múltiplos e comerciais terem reportado ao COSIF em pelo menos algum mês de 2004, apenas 77 apresentaram todos dados necessários para o cômputo da nova medida de produto bancário e das quantidades e preços de insumos a serem usados na estimação de fronteiras de produção e custo.

Mais especificamente, foram retirados da amostra bancos que reportaram possuir apenas um (ou nenhum) funcionário durante algum mês, ou que não reportaram dados sobre saldos de depósitos e/ou empréstimos concedidos. Também foram retirados alguns bancos que reportaram valores negativos como despesas administrativas ou gastos com trabalhos.

Considerou-se que, para a produção da nova medida de produto bancário descrita na Seção 2.3, os bancos utilizam três insumos:¹⁶

- X_1 : trabalho, medido em número de funcionários, sendo que o preço deste insumo, W_1 , foi definido como a razão entre despesas com trabalho e número de funcionários.
- X_2 : despesas administrativas, medidas em reais. Seu preço W_2 foi calculado como em Costa e Nakane (2004), através da razão entre despesas administrativas e ativo circulante.
- X_3 ou K : capital físico, medido em reais. O preço do capital fixo não foi calculado, dado que, na estimação da fronteira de custo, este é considerado como um insumo quasi-fixo.

Note-se que nenhum insumo financeiro foi utilizado, devido aos motivos discutidos no Capítulo 2.¹⁷ O Quadro 3.1 apresenta estatísticas descritivas das variáveis calculadas. Os

¹⁶ O Apêndice 1 apresenta uma descrição mais detalhada do cálculo destas variáveis.

¹⁷ Note-se, entretanto, que Wang (2003a, 2003b, 2003c, 2003d), ao aplicar sua nova medida do produto bancário, considera “*cash reserves*” como insumo da produção bancária, sem contudo apresentar motivação teórica para tanto.

valores são constantes de dezembro de 2004, corrigidos pelo Índice Geral de Preços – Disponibilidade Interna (IGP-DI) da Fundação Getúlio Vargas.

Quadro 3.1: Sumário da Amostra (em R\$ constantes de Dezembro/2004)

	Obs.	Média	Desvio Padrão	Valor Mínimo	Valor Máximo
Prod. Implícito Crédito (1)	924	808.175	1.832.438	-258.441	15.313.410
Prod. Implícito Depósitos (2)	924	31.008.714	118.561.386	-3.881.315	924.391.896
Produto Explícito (3)	924	26.573.524	82.408.462	0	533.162.910
Produto Total (1+2+3)	924	58.390.413	199.360.750	1.866	1.456.417.363
Capital Físico	924	116.245.970	348.213.353	17.535	2.442.031.264
Desp. Administrativas	924	12.289.234	36.928.039	2.331	316.831.104
Trabalho*	924	5.718	18.042	6	106.019
Preço – Trabalho	924	7.888	5.225	19	63.380
Preço – Desp. Adm.	924	129	125	3	725
Custo Variável	924	44.077.524	131.507.077	90.960	1.028.599.501

* - medido em número de funcionários.

Algumas inferências podem ser feitas sobre os resultados apresentados no Quadro 3.1. Primeiramente, o desvio-padrão do produto calculado é elevado quando comparado à sua média, indicando a presença de bancos relativamente grandes e pequenos. A próxima seção apontará que este é um fator relevante para a escolha da forma funcional da fronteira de produto (ou custo) a ser escolhida.

Em segundo lugar, note-se que ambos os itens do produto implícito apresentam valor mínimo negativo. Isto pode ser interpretado como a ocorrência de subsídios cruzados entre os componentes de produto de um banco.

A comparação entre a nova medida de produto bancário e a mais comumente usada (definida como a soma do saldo de operações de crédito e títulos e valores mobiliários¹⁸) aponta para uma proximidade entre as duas medidas. A correlação entre elas é de 0.98, bem acima daquela observada por Humphrey (1992) ao comparar a variável de estoque tradicional com o número de operações de crédito e depósitos realizados – 0.60.¹⁹

¹⁸ Contas COSIF 1.6.0.00.00-1 e 1.3.0.00.00-4, respectivamente.

¹⁹ Ver Seção 2.1.

3.2 Metodologia

Primeiramente, é preciso definir alguns conceitos relacionados à análise de eficiência. Coelli, Rao e Battese (1998) apontam inicialmente para dois componentes: a *eficiência técnica*, que consiste na capacidade da firma obter o máximo de produto(s) de um dado vetor de insumos, e a *eficiência alocativa*, que reflete a capacidade da firma em utilizar os insumos em proporções ótimas. Por sua vez, a combinação das duas forma a *eficiência econômica*.²⁰

O Quadro 3.2 ilustra graficamente estes conceitos para uma firma que usa dois insumos (x_1 e x_2) para produzir um único produto y sob retornos constantes de escala. A fronteira ótima é representada pela isoquanta SS' , enquanto AA' ilustra o preço relativo dos insumos.

Quadro 3.2: Representação Gráfica da Eficiência Técnica e Alocativa

Fonte: Coelli, Rao e Battese (1998).

Note-se que a ineficiência técnica de uma firma operando no ponto P é medida pela distância QP , que iguala o montante pelo qual todos insumos poderiam ser reduzidos sem que haja uma redução na quantidade produzida. Uma medida comum para eficiência técnica é a razão OQ/OP (Coelli, Battese e Rao, 1998). Note-se que esta razão varia entre zero e um, sendo que uma firma com medida igual a um é plenamente eficiente.

²⁰ Kumbhakar e Lovell (2000) usam o termo *eficiência de custo* (“cost efficiency”). Outros estudos que estimam fronteiras estocásticas para a indústria bancária brasileira (Nakane, 1999; Silva e Jorge Neto, 2002) usam o termo eficiência-X.

Já a eficiência alocativa pode ser medida pela razão OR/OQ , dado que a distância QR mede a redução de custos que ocorreria caso a firma produzisse no ponto alocativamente eficiente Q' .

Já a eficiência de escala pode ser mensurada ao analisar se uma firma opera com retornos crescentes ou de crescentes de escala. As fórmulas das medidas de eficiência técnica, econômica e de escala são dadas abaixo, pelas equações (10)-(13).

Para analisar a eficiência dos bancos brasileiros, os seguintes modelos são estimados:

$$\ln Y_{it} = \ln Y(X_{it}, \psi) + v_{it} - u_{it} \quad (4)$$

$$\ln C_{it} = \ln C(Y_{it}, W_{it}, K_{it}, \psi) + v_{it} + u_{it} \quad (5)$$

A equação (4) especifica uma fronteira de produção para a firma i no período t , onde o produto Y_{it} é produzido pela combinação do vetor de insumos X_{it} . Já a equação (5) especifica uma fronteira de custo, onde C_{it} representa o custo variável, e W_{it} e K_{it} são os vetores de preços de insumos variáveis e quantidades de insumos quasi-fixos, respectivamente.

Note-se que Y_{it} consiste de um único produto, formado pela soma dos dois produtos implícitos e o produto explícito calculado na seção anterior. A escolha por não usar estas três variáveis como produtos distintos segue Wang (2003c) e é feita para facilitar a abordagem via fronteira de produção, assim como evitar problemas relacionados a logaritmos de números negativos.

Em ambos casos, há um vetor de parâmetros β a ser estimado e v_{it} representa um termo aleatório com distribuição i.i.d. $N(0, \sigma_v^2)$. Por sua vez, u_{it} é um termo aleatório com distribuição i.i.d. $N^+(\mu, \sigma_u^2)^{21}$ que capta a ineficiência da firma i no período t .

Note-se que, no caso da fronteira de produção, u_{it} mensura a ineficiência técnica, uma vez que ela mede o quanto a firma i poderia estar produzir com o vetor de quantidades de insumos X_{it} caso fosse plenamente eficiente. Por outro lado, no caso da fronteira de custo, u_{it} mensura a ineficiência econômica (i.e., técnica e alocativa), uma vez que se mede apenas em quanto a

²¹ Isto é, trata-se de uma distribuição normal truncada no ponto zero com média μ .

firma i poderia reduzir seu custo sem diminuir sua produção, sem diferenciar se a fonte desta ineficiência é o uso de proporção sub-ótima de insumos ou ineficiência técnica.

Deste modo, é importante ressaltar que não se deve esperar uma proximidade entre as medidas de ineficiência estimadas através das fronteiras de custo e produção. O modelo incorpora variáveis determinantes da ineficiência ao especificar uma equação para μ :

$$\mu = \theta_0 + \sum_{j=1}^{10} \theta_j D_{jit} \quad (6)$$

Onde D_{jit} é a j -ésima variável que afeta a média da ineficiência de firma i no período t e θ_j são parâmetros a serem estimados. As variáveis utilizadas são apresentadas na Seção 3.5.

A equação (6) é estimada conjuntamente com as equações (4) ou (5) em uma única etapa, através do método da máxima verossimilhança.²² Neste estudo, foi utilizado o *software* Stata®, Versão 9.1, para todas estimações.

A metodologia acima já foi utilizada por Nakane (1999) e Silva e Jorge Neto (2002) para estimar fronteiras de custo *translog* para a indústria bancária brasileira nos períodos 1990-97 e 1995-99, respectivamente.

Diferentemente, neste estudo, optou-se por utilizar a forma flexível de Fourier, originalmente desenvolvida em Gallant (1981,1982). Esta forma baseia-se no fato de que uma função pode ser aproximada por uma combinação linear de termos trigonométricos denominada expansão de Fourier. Isto ocorre pois as funções seno e co-seno são mutualmente ortogonais no espaço $[0, 2\pi]$, permitindo a aproximação de uma função desconhecida.

Cumprе notar que a forma de Fourier é usualmente aplicada de modo a englobar a especificação *translog* (que, por sua vez, engloba a especificação Cobb-Douglas), permitindo que uma seja testada contra a outra. Berger, Leusner e Mingo (1997) argumentam que a aproximação polinomial proporcionada pela *translog* não é adequada para dados bancários onde exista grande variância em relação à escala média.

²² A derivação da função de máxima verossimilhança do modelo, originalmente feita em Battese e Coelli (1995), pode ser encontrada em StataCorp (2005).

Os autores também apontam que:

“The translog imposes a U-shape on the average cost curve that is symmetric in logs around the bottom of the U, and has no other inflection points. This may distort scale efficiency measurement, since it forces any scaled diseconomies for larger entities to take the same shape as scale economies for smaller entities” (Berger, Leusner e Mingo, 1997).

A forma flexível de Fourier também já foi usada na estimação de fronteiras estocásticas para o setor bancário em Mitchell e Onvural (1996)²³; Berger e Mester (1997); e Altunbas et al. (2001). Contudo, apenas este último utiliza a forma de Fourier na estimação de uma fronteira estocástica por máxima verossimilhança, sendo o que mais se aproxima metodologicamente deste trabalho.

Assim, a forma funcional da fronteira de produção a ser estimada é dada por uma especificação *translog* adicionada de uma expansão de Fourier de segunda ordem:²⁴

$$\ln Y = \beta_0 + \sum_{i=1}^3 \beta_i \ln X + \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \beta_{ij} \ln X_{ij} \ln X_{ij} + \sum_{i=1}^3 [a_i \sin(z_i) + b_i \cos(z_i)] + \sum_{i=1}^3 \sum_{j=1}^3 [a_{ij} \sin(z_i + z_j) + b_{ij} \cos(z_i + z_j)] + \sum_{i=1}^3 \tau_i T \ln X_i + \sum_{t=1}^{11} \lambda_t + v - u \quad (7)$$

Onde os β s são os parâmetros de uma forma funcional *translog* e λ_t são variáveis *dummy* para cada período. Note-se que, assim como em Altunbas et al. (2001), são incluídos parâmetros τ relacionados às interações entre uma variável de tendência (T) e as quantidades de insumos X .

Por sua vez, os a s e b s são parâmetros relacionados à expansão de Fourier, sendo que Z_i são valores ajustados das quantidades de insumos X_i , satisfazendo que Z_i pertença ao intervalo $[0, 2\pi]$. Os demais termos estão definidos na equação (4).

²³ Este estudo utiliza a forma de Fourier na estimação de uma função de custo, e não de uma fronteira.

²⁴ Para facilitar a notação, os subscritos indexadores de firma e período (i e t) são omitidos das equações (7) e (9).

A construção dos Z s segue a metodologia de Berger, Leusner e Mingo (1997), que define:²⁵

$$z_i = 0.2\pi + \frac{1.8\pi - 0.2\pi}{\max(\ln X_i) - \min(\ln X_i)} [\ln X_i - \min(\ln X_i)] \quad (8)$$

Já a forma funcional da fronteira de custo é dada por:

$$\begin{aligned} \ln C = & \beta_0 + \sum_{i=1}^4 \beta_i \ln U + \frac{1}{2} \sum_{i=1}^4 \sum_{j=1}^4 \beta_{ij} \ln U_{ij} \ln U_{ij} + \sum_{i=1}^4 [a_i \sin(z_i) + b_i \cos(z_i)] + \\ & \sum_{i=1}^4 \sum_{j=1}^4 [a_{ij} \sin(z_i + z_j) + b_{ij} \cos(z_i + z_j)] + \sum_{i=1}^4 \tau_i T \ln U_i + \sum_{t=1}^{11} \lambda_t + v + u \end{aligned} \quad (9)$$

Onde $U_i = \{W_1, W_2, K, Y\}$. Os Z s são as quantidades ajustadas destas quatro variáveis, calculadas de maneira análoga à feita no caso da fronteira de produção.

Note-se que a propriedade de homogeneidade linear da função custo nos preços de insumos não foi imposta. Devido à dificuldade de impor esta propriedade como restrição dos parâmetros de termos trigonométricos, alguns trabalhos²⁶ utilizam apenas senos e co-senos das quantidades ajustadas de produto, insumos quasi-fixos e outras variáveis incluídas na função, deixando que o efeito dos preços de insumos no custo seja totalmente definido pelos parâmetros de uma função *translog*, onde a homogeneidade pode ser facilmente imposta como uma restrição dos parâmetros.

Por outro lado, trabalhos como Mitchell e Onvural (1996) e Berger e Mester (1997) incluem termos trigonométricos dos preços de insumos, com base no argumento de que mesmo sem a imposição de homogeneidade linear, a forma flexível de Fourier é capaz de aproximar uma função custo que satisfaça esta propriedade (Gallant, 1982).

Uma vez estimada as fronteiras de produção, é possível obter a medida de eficiência técnica para cada firma i no período t através de:

²⁵ Note-se que tal fórmula faz com que os Z s pertençam ao intervalo $[0.1*2\pi, 0.9*2\pi]$, evitando problemas de aproximação nos limites do intervalo $[0, 2\pi]$, conforme discutido em Gallant (1981).

²⁶ Berger, Leusner e Mingo (1997); e Altunbas et al. (2001), por exemplo.

$$Eftec_{it} = E[\exp(-u_{it}) | \ln Y_{it} - \ln \hat{Y}_{it}] \quad (10)$$

Onde $\hat{Y}_{it}$ é a quantidade de produto prevista pela fronteira estimada. Esta medida de ineficiência é proposta por Battese e Coelli (1988), e usa o fato de que, omitindo o termo de ruído aleatório v_{it} , temos $\exp(-u_{it}) = Y_{it}/Y(X_{it}, \beta)$. De tal modo que $0 \leq Eftec \leq 1$ mensura qual a porcentagem do potencial produtivo de uma firma que é de fato utilizada.²⁷ Por sua vez, a medida de economias de escala é dada por.

$$ES_{PROD} = \sum_{j=1}^3 \frac{\partial \ln \hat{Y}_{it}}{\partial \ln X_{jit}} \quad (11)$$

Onde $ES_{PROD} > 1$ indica a presença de retornos crescentes de escala e $ES_{PROD} < 1$ a presença de retornos decrescentes de escala, enquanto $ES_{PROD} = 1$ aponta para retornos constantes. Note-se que as economias de escala são medidas assumindo que a firma se encontra sobre a fronteira eficiente.

Já no caso da fronteira e custo, as medidas de eficiência econômica e de economias de escala são dadas pelas equações (12) e (13).

$$Efecon_{it} = 1 / E[\exp(u_{it}) | \ln C_{it} - \ln \hat{C}_{it}] \quad (12)$$

$$ES_{CUSTO} = 1 / \frac{\partial \ln C_{it}}{\partial \ln Y_{it}} \quad (13)$$

Onde $\hat{C}_{it}$ é o custo variável previsto pela fronteira estimada. A interpretação de ES_{CUSTO} é a mesma de ES_{PROD} . De modo análogo ao caso de *Eftec*, *Efecon* mensura qual a porcentagem do custo da firma é usada de maneira eficiente.²⁸

²⁷ A definição da ineficiência técnica como a razão entre a quantidade produzida e a quantidade que seria produzida por uma firma plenamente eficiente (situada na fronteira) é denominada “medida de Debreu-Farrell” (Coelli, Rao e Battese, 1998).

²⁸ A derivação da fórmula das medidas de ineficiência está disponível em StataCorp (2005) e no artigo original de Battese e Coelli (1988). Altunbas et al. (2001) apresenta a fórmula da derivada de uma forma de Fourier.

Por fim, é necessário apontar brevemente para algumas das hipóteses implícitas desta metodologia.²⁹ Primeiramente, note-se que o modelo assume a homocedasticidade tanto dos termos de ineficiência u_{it} como do choque aleatório v_{it} .

Em segundo lugar, a hipótese de que os termos de ineficiência, u_{it} , são independentes e identicamente distribuídos com média determinada pelas variáveis D_{jit} implica na independência de u_{it} em relação às variáveis da fronteira de produção e dos choques aleatórios captados por v_{it} . Esta hipótese é especialmente restritiva quando consideramos que os bancos podem considerar seu grau de ineficiência na decisão da escolha da quantidade de insumos (no caso da função de produção) e produto (no caso da fronteira de custo). Caso isto ocorra, problemas de endogeneidade podem ocorrer com o modelo.

Em terceiro lugar, também se assume que toda dependência no tempo da ineficiência de certo banco é dada pela variação dos determinantes de μ . Isto é, o modelo impõe uma estrutura rígida para a autocorrelação do termo de ineficiência de um banco ao longo do tempo, enquanto assume a inexistência de qualquer forma de correlação entre os choques aleatórios dos indivíduos da amostra.

As próximas quatro seções apresentam o resultado da aplicação da metodologia descrita acima aos dados discutidos na Seção 3.1

3.3 Resultados: Fronteira de Produção

As equações (4) e (6) foram conjuntamente estimadas com base na amostra descrita na Seção 3.1. Observou-se a presença de vários coeficientes insignificantes na fronteira de produção, muito provavelmente decorrente da multicolinearidade entre as 41 variáveis que constituem a fronteira.

Deste modo, optou-se por uma abordagem “*general to specific*” para a obtenção de um modelo mais parcimonioso, que consiste em repetidamente excluir a variável cujo coeficiente apresentava maior p-valor e re-estimar o modelo até o ponto em que todos os coeficientes são significantes. Uma vez definidas as variáveis a serem incluídas na fronteira de produção, a

²⁹ Kumbhakar e Lovell (2000) apresentam uma discussão mais aprofundada destas questões.

mesma abordagem foi utilizada para excluir as variáveis determinantes da ineficiência que se mostraram insignificantes.

Os coeficientes da estimação do modelo são apresentados no Quadro 3.3. O Apêndice 2 reporta resultados do modelo “irrestrito” e de testes de hipótese validando as restrições impostas.

Quadro 3.3: Parâmetros da Fronteira de Produção

Erro-				Erro-			
Variável	Coefficiente	Padrão	P-Valor	Variável	Coefficiente	Padrão	P-Valor
$\ln(X_1)$	10.693	1.438	0.000	$\text{sen}(2Z_2)$	-1.922	0.264	0.000
$\ln(X_2)$	-11.971	1.470	0.000	$\text{sen}(Z_2+Z_3)$	0.967	0.182	0.000
$\ln(X_1)^2$	-0.690	0.175	0.000	$\text{sen}(Z_1+Z_2)$	1.820	0.311	0.000
$\ln(X_3)^2$	0.052	0.009	0.000	λ_1	-0.342	0.079	0.000
$\ln(X_2)^2$	1.006	0.114	0.000	λ_2	-0.245	0.077	0.001
$\ln(X_1)\ln(X_3)$	-0.068	0.018	0.000	λ_3	-0.310	0.077	0.000
$\ln(X_1)\ln(X_2)$	-0.258	0.045	0.000	λ_4	-0.540	0.077	0.000
$\cos(Z_1)$	6.725	1.214	0.000	λ_5	-0.397	0.079	0.000
$\cos(Z_2)$	-8.890	0.810	0.000	λ_6	-0.355	0.078	0.000
$\cos(2Z_1)$	1.670	0.206	0.000	λ_7	-0.465	0.078	0.000
$\cos(Z_1+Z_3)$	-0.361	0.097	0.000	λ_8	-0.396	0.078	0.000
$\cos(Z_1+Z_2)$	-2.323	0.216	0.000	λ_9	-0.374	0.078	0.000
$\text{sen}(Z_1)$	3.394	0.319	0.000	λ_{10}	-0.373	0.077	0.000
$\text{sen}(Z_3)$	1.405	0.279	0.000	λ_{11}	-0.326	0.077	0.000
$\text{sen}(Z_2)$	-2.053	0.546	0.000	constante	52.583	9.760	0.000
$\text{sen}(2Z_1)$	-0.839	0.158	0.000				

O Quadro 3.4 apresenta as elasticidades-insumo da fronteira de produção, visando checar se a fronteira estimada apresenta os resultados esperados pela teoria econômica. Note-se que, devido à presença de termos quadráticos e trigonométricos, estas elasticidades são funções das próprias quantidades de insumo. Assim, os valores apresentados foram computados para o valor médio da amostra.

Quadro 3.4: Elasticidades da Fronteira de Produção

Insumo	Elasticidade	Estat. Wald	P-Valor
Trabalho	0.320	87.90	0.000
Desp. Admin.	0.425	117.14	0.000
Capital Físico	0.140	34.16	0.000

O Quadro 3.4 indica que as elasticidades-insumo têm o valor esperado pela teoria econômica. Ademais, o teste de Wald rejeita fortemente a hipótese de que as elasticidades estimadas são iguais à zero.

Também foram computadas as elasticidades-insumo, assim como sua estatística de Wald, para todos 924 pontos da amostra. Aproximadamente 58% das unidades banco-mês apresentaram elasticidades do capital e das despesas administrativas positivas e estatisticamente significantes ao nível de 5%, enquanto este número cai para 43% no caso da elasticidade-trabalho.³⁰

Com o intuito de checar se a especificação utilizada é adequada, o Quadro 3.5 reporta o resultado de uma série de testes Wald de restrições sobre o modelo.³¹

Quadro 3.5: Testes de Especificação da Fronteira de Produção

Hip. Nula	Estat. Wald	P-Valor
1) Todos parâmetros do modelo = 0	7179.71	0.000
2) $a_i=b_i=a_{ij}=b_{ij}=0$, para todo i,j	278.14	0.000
3) $\beta_i=\beta_{ij}=a_i=b_i=a_{ij}=b_{ij}=0$, para todo i,j	477.39	0.000

O fato da hipótese nula (1) ser rejeitada indica para a qualidade geral do modelo. Já as hipóteses nulas (2) e (3) testam, respectivamente, se uma especificação *translog* ou Cobb-Douglas seria mais adequada. Ambas são fortemente rejeitadas em favor da forma funcional de Fourier.

³⁰ 25%, 14% e 11% dos pontos amostrais apresentaram, respectivamente, elasticidades-trabalho, -capital e – despesas administrativas, negativas e significantes a 5%.

³¹ Testes de hipótese acerca o termo de ineficiência v_{it} estão disponíveis na Seção 3.5.

3.4 Resultados: Fronteira de Custo

Assim como no caso da fronteira de produção, as equações (5) e (6) foram conjuntamente estimadas com base na amostra descrita na Seção 3.1. A mesma estratégia “*general to specific*” foi adotada. Contudo, optou-se por manter algumas variáveis não significantes que se mostraram necessárias para que o algoritmo de maximização da função verossimilhança do modelo atingisse convergência.

Quadro 3.6: Parâmetros da Fronteira de Custo

Erro-				Erro-			
Variável	Coefficiente	Padrão	P-Valor	Variável	Coefficiente	Padrão	P-Valor
$\ln(W_1)$	3.926	1.453	0.007	$\text{sen}(Z_3)$	0.382	0.119	0.001
$\ln(K)$	5.043	0.785	0.000	$\text{sen}(Z_4)$	-1.761	0.170	0.000
$\ln(Y)$	-4.349	0.943	0.000	$\text{sen}(2Z_1)$	-0.162	0.061	0.008
$\ln(W_2)$	-5.335	1.720	0.002	$\text{sen}(2Z_2)$	-0.224	0.051	0.000
$\ln(W_1)^2$	-0.297	0.105	0.004	$\text{sen}(2Z_3)$	-0.453	0.056	0.000
$\ln(K)^2$	-0.248	0.048	0.000	$\text{sen}(2Z_4)$	-0.837	0.088	0.000
$\ln(Y)^2$	0.445	0.066	0.000	$\text{sen}(Z_1+Z_2)$	0.173	0.018	0.000
$\ln(W_2)^2$	1.035	0.366	0.005	$\text{sen}(Z_2+Z_4)$	0.016	0.012	0.161
$\ln(K)\ln(Y)$	-0.064	0.013	0.000	$\text{sen}(Z_1+Z_3)$	-0.043	0.013	0.001
$\ln(W_2)\ln(K)$	-0.020	0.010	0.050	$\text{sen}(Z_1+Z_4)$	0.057	0.014	0.000
$\ln(W_1)\ln(W_2)$	0.067	0.028	0.016	$\text{sen}(Z_3+Z_4)$	0.769	0.110	0.000
$\cos(Z_1)$	3.063	0.888	0.001	$\ln(W_2)T$	0.004	0.002	0.047
$\cos(Z_2)$	-1.954	0.685	0.004	λ_1	0.160	0.101	0.111
$\cos(Z_3)$	3.111	0.483	0.000	λ_2	0.180	0.093	0.053
$\cos(Z_4)$	-3.605	0.653	0.000	λ_3	0.160	0.084	0.057
$\cos(2Z_1)$	0.526	0.139	0.000	λ_4	0.135	0.076	0.076
$\cos(2Z_2)$	-0.152	0.095	0.110	λ_5	0.119	0.069	0.085
$\cos(2Z_3)$	0.552	0.073	0.000	λ_6	0.089	0.061	0.141
$\cos(2Z_4)$	-0.612	0.089	0.000	λ_7	0.070	0.054	0.194
$\cos(Z_2+Z_4)$	-0.028	0.011	0.008	λ_8	0.071	0.047	0.126
$\cos(Z_1+Z_3)$	0.033	0.013	0.012	λ_9	0.069	0.041	0.089
$\cos(Z_1+Z_4)$	-0.068	0.015	0.000	λ_{10}	0.056	0.036	0.115
$\text{sen}(Z_1)$	-0.635	0.180	0.000	λ_{11}	0.088	0.030	0.004
$\text{sen}(Z_2)$	-0.611	0.157	0.000	constante	-18.487	11.009	0.093

Os coeficientes da estimação do modelo são apresentados no Quadro 3.6. O Apêndice 2 reporta resultados do modelo “geral” e de testes de hipótese validando as restrições impostas.

As variáveis Z_1 , Z_2 , Z_3 e Z_4 (cuja construção foi discutida na seção anterior) se referem às quantidades ajustadas de W_1 , W_2 , K e Y , respectivamente.

Cumprir testar se a fronteira de custo estimada atende as propriedades previstas pela teoria para uma função custo: i) ser não-decrescente em W_i e Y e ii) ser côncava em W_i .³² A propriedade (i) pode ser facilmente analisada através das elasticidades-preço e elasticidades-produto da fronteira de custo. Para testar a propriedade (ii), utiliza-se o fato de que o Hessiano (matriz de segundas derivadas) de uma função côncava é uma matriz semi-definida negativa, sendo que uma matriz simétrica é semi-definida negativa se, e somente se, todos seus autovalores forem não-positivos.

O Quadro 3.7 apresenta os testes das propriedades (i) e (ii). Como as elasticidades e autovalores são funções das variáveis da fronteira de custo, são reportados os valores dos testes computados no ponto médio da amostra. Ademais, os autovalores são funções não-lineares dos parâmetros estimados, sendo que o “método-delta” foi usado para calcular a variância usada para computar suas estatísticas de teste.

Quadro 3.7: Testes de Propriedades da Fronteira de Custo

	Valor	Estat.-Wald	P-Valor
Elast.-Preço (Trabalho)	0.230	188.73	0.000
Elast.-Preço (Desp. Adm.)	5.093	8.84	0.003
Elast.-Produto	1.407	122.10	0.000
Autovalor ₁	0.041	0.41	0.523
Autovalor ₂	-0.718	9.80	0.002

O fato das três elasticidades indicadas no Quadro 3.7 serem positivas e estatisticamente significantes indica que a propriedade (i) é atendida. Ademais, ao computar as elasticidades para cada indivíduo (i.e., unidade banco-mês) da amostra, são observadas elasticidades do

³² Outra propriedade das funções custo é a homogeneidade linear em W_i , que é usualmente imposta na estimação de funções custo através de restrições nos parâmetros. Contudo, conforme discutido na Seção 3.2, esta imposição não foi feita neste trabalho.

preço do trabalho e das despesas administrativas positivas em, respectivamente, 97.6% e 100% dos casos, enquanto 99.9% da amostra apresenta elasticidade-produto positiva.

Note-se também que um dos autovalores não é estatisticamente diferente de zero, enquanto o outro é significantemente negativo, apontando para o fato de que a propriedade (ii) também é atendida.

Assim como feito na Seção 3.3 para o caso da fronteira de produção, o Quadro 3.8 apresenta testes de especificação para a fronteira de custo, deixando os testes envolvendo o termo de ineficiência para a próxima seção.

Quadro 3.8: Testes de Especificação da Fronteira de Custo

Hip. Nula	Estat. Wald	P-Valor
1) Todos parâmetros do modelo = 0	36064.43	0.000
2) $a_i=b_i=a_{ij}=b_{ij}=0$, para todo i,j	724.65	0.000
3) $\beta_i=\beta_{ij}=a_i=b_i=a_{ij}=b_{ij}=0$, para todo i,j	1030.80	0.000

Assim como no caso da fronteira de produção (Seção 3.3), os testes de especificação apontam para a significância conjunta do modelo e adequação da forma funcional de Fourier, rejeitando a hipótese nula de que o modelo estimado seja uma *translog* ou Cobb-Douglas.

3.5 Resultados: Determinantes da Ineficiência Técnica e Econômica

Esta seção discute a estimação do termo u_{it} , que mede a ineficiência técnica (no caso da fronteira de produção) ou econômica (no caso da fronteira de custo). Cumpre relembrar que este termo apresenta distribuição normal truncada $N^+(\mu, \sigma_u^2)$, onde μ é dado pela combinação linear das seguintes variáveis determinantes da eficiência:³³

- *Ativo*: logaritmo natural do total de ativos do banco.
- *Ativo*²: quadrado da variável *Ativo*.
- *Plativo*: razão entre patrimônio líquido e total de ativos.
- *Empativo*: razão entre saldo de fundos emprestados e total de ativos.

³³ O Apêndice 2 discute a construção destas variáveis em maior detalhe.

- *Depav*: razão entre saldo de depósitos à vista e total de ativos.
- *Nonperform*: razão entre o saldo de créditos com mais de 90 dias de atraso e o saldo total de crédito.
- *Público*: variável *dummy* assumindo valor um para bancos públicos.
- *Cestrang*: variável *dummy* assumindo valor um para bancos com controle estrangeiro.
- *Pestrang*: variável *dummy* assumindo valor um para bancos com participação acionária de estrangeiros, porém controlados por grupos nacionais.

Estas variáveis são também usadas em Nakane (1999) e/ou Silva e Jorge Neto (2002), o que possibilita comparações entre os resultados obtidos.

As variáveis *ativo* e *ativo*² são incluídas para captar uma possível relação (não-linear) entre tamanho do banco e eficiência. A variável *plativo* capta o efeito de uma maior participação de capital próprio no financiamento do banco em sua eficiência. Segundo Silva e Jorge Neto (2002), a teoria do risco moral sugere que esta variável esteja negativamente relacionada à ineficiência, pois, em caso de insolvência, a maior presença de capital próprio levaria a maiores prejuízos para os acionistas.

As variáveis *empativo* e *depav* testam se a fonte de financiamento do banco afeta sua eficiência, enquanto a variável *nonperform* mensura o impacto da qualidade do crédito sobre a ineficiência. Por fim, *publico*, *pestrang* e *cestrang* visam medir as diferenças de eficiência entre bancos com diferente estrutura de propriedade.

O Quadro 3.9 apresenta os parâmetros estimados do modelo de ineficiência para a fronteira de produção. Duas especificações são reportadas, a irrestrita, com todas variáveis determinantes da ineficiência, e a restrita, que exclui as variáveis insignificantes. Note-se que ambas tomam como base a especificação da fronteira que consta no Quadro 3.3.

Por sua vez, o Quadro 3.10 apresenta os resultados acerca da ineficiência econômica, obtida através da fronteira de custo. Cumpre notar que uma variável (*ativo*) não significativa foi incluída no modelo restrito, pois o conjunto de variáveis relacionadas ao tamanho do banco (*ativo* e *ativo*²) é estatisticamente significativo. Tanto o modelo irrestrito como o restrito toma como base a especificação da fronteira que consta no Quadro 3.6.

Quadro 3.9: Parâmetros da Ineficiência Técnica (Fronteira de Produção)

	Modelo Irrestrito			Modelo Restrito		
	Coef.	Erro-Padrão	P-Valor	Coef.	Erro-Padrão	P-Valor
<i>constante</i>	-15.467	10.995	0.160	-20.376	7.998	0.011
<i>ativo</i>	2.480	1.062	0.020	2.927	0.799	0.000
<i>ativo</i> ²	-0.081	0.026	0.002	-0.091	0.020	0.000
<i>empativo</i>	1.358	0.194	0.000	1.370	0.178	0.000
<i>publico</i>	-1.426	0.305	0.000	-1.715	0.244	0.000
<i>pestrang</i>	-3.754	1.447	0.009	-3.803	1.429	0.008
<i>plativo</i>	-0.008	0.455	0.985	-	-	-
<i>depav</i>	-1.974	1.632	0.227	-	-	-
<i>nonperform</i>	-0.332	0.175	0.057	-	-	-
<i>cestrang</i>	0.113	0.098	0.249	-	-	-
σ_u^2	0.666	0.052	*	0.749	0.053	*
γ	0.908	0.014	*	0.912	0.013	*

Legenda: * - significativo ao nível de 5%.

Quadro 3.10: Parâmetros da Ineficiência Econômica (Fronteira de Custo)

	Modelo Irrestrito			Modelo Restrito		
	Coef.	Erro-Padrão	P-Valor	Coef.	Erro-Padrão	P-Valor
<i>constante</i>	-2.007	3.986	0.615	-1.657	3.990	0.678
<i>ativo</i>	0.476	0.325	0.143	0.446	0.325	0.171
<i>ativo</i> ²	-0.017	0.007	0.013	-0.016	0.007	0.018
<i>plativo</i>	0.907	0.096	0.000	0.847	0.085	0.000
<i>depav</i>	2.048	0.245	0.000	2.014	0.243	0.000
<i>publico</i>	0.414	0.039	0.000	0.422	0.038	0.000
<i>pestrang</i>	-0.532	0.092	0.000	-0.530	0.094	0.000
<i>nonperform</i>	-0.116	0.039	0.003	-0.123	0.038	0.001
<i>empativo</i>	0.048	0.049	0.332	-	-	-
<i>cestrang</i>	-0.030	0.027	0.261	-	-	-
σ_u^2	0.041	0.003	*	0.041	0.003	*
γ	0.877	0.035	*	0.877	0.035	*

Legenda: * - significativo ao nível de 5%.

Primeiramente, é importante analisar o termo de variância σ_u^2 . Em ambos casos, ele é estatisticamente significativo. Ademais, a estatística $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$ ³⁴ indica que a maior parte dos desvios em relação a fronteira é devido à ineficiência, e não ao ruído aleatório (v_{it}).

É importante ressaltar que, conforme discutido na Seção 3.2, a fronteira de produção permite observar a eficiência técnica, enquanto a fronteira de custo permite observar a eficiência econômica. Assim, a diferença entre os resultados dos Quadros 3.9 e 3.10 é esperada, e pode inclusive ser interpretada como um indicador de que a eficiência alocativa é um fator relevante na estimação de fronteiras de custo para o setor bancário brasileiro.

Testes de Wald apontaram para a significância conjunta, ao nível de 1%, das variáveis incluídas nos quatro modelos dispostos nos Quadros 3.9 e 3.10.

Ambos os casos apontam para uma relação com forma de “U” entre tamanho e eficiência. O mesmo resultado foi encontrado por Nakane (1999) e Silva e Jorge Neto (2002). Outro resultado comum entre as fronteiras de produção e custo é o impacto positivo da participação estrangeira (mantendo o controle por um grupo nacional) na eficiência.

Os resultados da fronteira de produção apontam para maior eficiência dos bancos públicos, enquanto os da fronteira de custo indicam o oposto. Uma possível interpretação para este resultado seria que, apesar de possuírem maior eficiência técnica, os bancos públicos sofrem de acentuada ineficiência alocativa nos insumos, que poderia ser explicada pela distinção entre funcionalismo público e privado que é feita pela legislação trabalhista brasileira. Note-se também que uma menor eficiência econômica dos bancos públicos corrobora o resultado de Silva e Jorge Neto (2002), em oposição ao obtido em Nakane (1999).

Os coeficientes da variável *empativo* indicam que bancos indicam que o maior uso de fundos emprestados eleva a ineficiência técnica. Segundo Silva e Jorge Neto (2002), este resultado é esperado pois o saldo de fundos emprestados contém tanto empréstimos de outras instituições financeiras e do Banco Central (operações de redesconto), usualmente feitas a taxas acima do nível de mercado, como repasses do Governo Federal, que podem não ser alocados de forma

³⁴ Ver Battese e Coelli (1995), assim como Kumbhakar e Lovell (2000), para maiores informações sobre esta estatística.

ótima pelo banco. Contudo, a estimação da fronteira de produção não indica para a relevância desta variável.

Os resultados da variável *depav* indicam que menor custo de captação dos depósitos à vista não se reverte em maior eficiência, estando inclusive associados à ineficiência econômica. Este resultado também é observado em Silva e Jorge Neto (2002), que sugere que o fato dos depósitos à vista estarem sujeitos a uma maior incidência de reserva compulsória pode levar à ineficiência.

A tese de que problemas de risco moral fazem com que bancos com maior participação de capital próprio sejam mais eficientes é rejeitada, conforme indicam os resultados dos coeficientes relacionados à variável *plativo*.

Por fim, a variável *nonperform* está associada a maior eficiência apenas no caso da fronteira de custo, sendo estatisticamente insignificante no caso da fronteira de produção. Este resultado pode ser interpretado pelo fato de que os bancos incorrem algum tipo de ineficiência alocativa ao tentarem evitar a exposição ao risco em suas operações de crédito de maneira excessivamente custosa.

3.6 Resultados: Medidas de Economias de Escala e Ineficiência

Esta seção apresenta os resultados das medidas de ineficiência, definidas nas equações (10)-(13), calculados com base nos modelos que excluem as variáveis não-significativas, i.e., utilizando os parâmetros reportados no Quadro 3.3 e do “modelo restrito” que consta no Quadro 3.9 (no caso da fronteira de produção) e no Quadro 3.6 e “modelo restrito” do Quadro 3.10 (no caso da fronteira de custo).

Apesar da amostra permitir o cômputo destas medidas para cada banco em cada período (mês), optou-se por reportar apenas os valores médios (ponderados pela quantidade de produto) para toda amostra (o ano de 2004), visando evitar problemas relacionados à sazonalidade.

O Quadro 3.11 apresenta estes resultados médios, divididos por quintis do valor do ativo total de cada banco, possibilitando observar a variação de economias de escala e eficiência para bancos de diferentes tamanhos.

Quadro 3.11: Medidas de Eficiência e Economias de Escala

Quartil de Ativo	Economias de Escala		Eficiência	
	Prod	Custo	Prod	Custo
1º Quintil	1.381	0.737	0.111	0.442
2º Quintil	0.466	0.718	0.283	0.476
3º Quintil	0.780	0.703	0.540	0.526
4º Quintil	0.937	0.749	0.808	0.449
5º Quintil	0.885	0.727	0.430	0.472
Total	0.885	0.727	0.430	0.472
Desvio-Padrão	0.915	0.105	0.319	0.246

Primeiramente, note-se que o baixo valor da medida de eficiência média da amostra. Os valores abaixo de 0.5 indicam que, em média, os bancos da amostra poderiam até mesmo dobrar sua produção (ou reduzir seus custos pela metade) sem utilizar mais insumos (ou reduzir sua produção).

Note-se também que estes valores estão abaixo dos encontrados por Nakane (1999) e Silva e Jorge Neto (2002), que estimaram a eficiência média de suas amostras como 0.62 e 0.86, respectivamente. Uma interpretação para isto é que as abordagens para calcular o produto bancário nestes trabalhos estariam superestimando esta variável.³⁵

Outra interpretação seria que estes estudos podem estar “confundindo” maior exposição ao risco com um maior grau de eficiência, o que é corroborado pelo fato de Nakane (1999), que inclui uma variável medindo exposição ao risco (valor dos créditos em atraso) na fronteira eficiente, encontrar uma eficiência média inferior àquela obtida por Silva e Jorge Neto (2002), que não apresentam controles para exposição ao risco em sua fronteira de custo.

³⁵ Note-se que, pelo fato da medida adotada neste trabalho ser uma variável de fluxo, enquanto a usualmente adotada constitui variável de estoque, que não se pode comparar a magnitude delas diretamente. De todo modo, cumpre lembrar que a amostra deste trabalho difere da usada em Nakane (1999) e Silva e Jorge Neto (2002), tanto pelos bancos incluídos como pelo período analisado.

O valor médio da medida de economia de escala obtido está próximo do observado por Silva e Jorge Neto (2002),³⁶ indicando que os bancos brasileiros operam abaixo de sua escala ótima, o que pode explicar a ocorrência relativamente freqüente de fusões e aquisições neste setor.

Por fim, cumpre comparar os resultados obtidos através da fronteira de produção e da fronteira de custo. Apesar dos valores médios observados serem próximos, as médias calculadas para bancos de diferentes portes apresentam acentuadas diferenças. À propósito, a correlação entre as economias de escala calculadas pelas duas fronteiras é de apenas 0.18, sendo que este valor cai para 0.08 no caso das medidas de ineficiência.

Note-se também que os valores de eficiência e economias de escala obtidos através da fronteira de produção apresentam variância superior às obtidas pela de custo.

³⁶ Os resultados acerca de economias de escala deste trabalho não serão comparados com o de Nakane (1999), que as mensura de maneira distinta.

4 CONCLUSÃO

Os pontos abaixo resumiam as principais conclusões deste trabalho:

1. As medidas de produção bancária comumente usadas na literatura nacional e internacional apresentam deficiências importantes, seja por constituírem variável de fluxo, por não definirem de maneira teoricamente embasada o papel dos insumos ou por relevarem o papel da exposição ao risco.
2. A abordagem dos ativos, que define que o valor de itens do ativo financeiro como produto e os depósitos bancários como insumo, é a mais utilizada na literatura brasileira. Contudo, as variações onde depósitos são considerados com produto ou são simplesmente relevados são igualmente freqüentes.
3. É possível, com base na metodologia de Wang (2003a, 2003b) e nos dados do COSIF, calcular uma medida de produção bancária que constitua uma variável de fluxo e incorpore de maneira teoricamente embasada o papel dos depósitos e da exposição ao risco.
4. Esta nova metodologia foi utilizada para estimar fronteiras estocásticas de produção e custo, visando estimar a eficiência técnica (e econômica) e economias de escala. O modelo utiliza uma forma funcional de Fourier para especificar a fronteira eficiente e ainda incorpora variáveis determinantes da eficiência.
5. Os dados usados constituem um painel de periodicidade mensal, incluindo 77 bancos múltiplos e comerciais durante o período de janeiro a dezembro de 2004.
6. Os resultados apontam para um elevado grau de ineficiência técnica, sendo que o banco médio da amostra poderia dobrar sua produção (ou reduzir seu custo pela metade) sem aumentar a quantidade de insumos utilizada (ou diminuir a quantidade produzida).

7. Também se conclui que os bancos brasileiros trabalham em ponto onde há retornos decrescentes de escala.
8. Apesar dos valores médios das medidas de economia de escala e eficiência calculadas pelas fronteiras de produção e custo serem próximos, estes apresentam baixa correlação e diferentes graus de variância. Ademais, o sinal e significância dos coeficientes de determinantes da ineficiência também dependem de como a tecnologia é representada.
9. O ponto acima indica que a escolha entre o uso de fronteiras de produção ou de custo acarreta diferenças nos resultados obtidos. Assim, deve-se considerar qual representação da tecnologia é a mais adequada para cada caso, não sendo possível indicar uma como mais adequada que a outra.
10. De todo modo, os resultados permitem realizar algumas recomendações para políticas públicas:
 - a. Os bancos brasileiros não incorporam totalmente suas economias de escala, o que ajuda a explicar o movimento recente de fusões e aquisições no setor.
 - b. O impacto do tamanho do banco na sua eficiência assume um formato de “U-invertido”, sendo que há um tamanho ótimo eficiente. Fusões que aproximem os bancos deste ponto ótimo também podem gerar eficiências.
 - c. A participação acionária de grupos estrangeiros eleva a eficiência bancária. Contudo, o controle externo da instituição financeira não parece ter impacto na eficiência.
 - d. Por um lado, a análise da fronteira de custo aponta para uma menor eficiência dos bancos públicos, como em Silva e Jorge Neto (2002). Por outro lado, a fronteira de produção indica o oposto, como em Nakane (1999).

REFERÊNCIAS

- ALTUNBAS, Yener, GARDENER, Edward P.M., MOLYNEUX, Philip & MOORE, Barry (2001), Efficiency in European banking. *European Economic Review*, **45**, 1931-1955.
- BANCO CENTRAL DO BRASIL (2006) *Economia Bancária e Crédito*. Banco Central do Brasil.
- BATTESE, George & COELLI, Tim J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, **38**, 387-399.
- BATTESE, George E. & COELLI, Tim J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, **20**, 325-332.
- BELAISCH, Agnes J. (2003), *Do Brazilian banks compete?* Washington: IMF Working Paper 03/113.
- BERGER, Allen N., LEUSNER, John H. & MINGO, John, J. (1997), The efficiency of bank branches. *Journal of Monetary Economics*, **40**, 141-162.
- BERGER, Allen N. & HUMPHREY, David B. (1992), Measurement and efficiency issues in commercial banking. In: GRILICHES, Zvi. (ed.), *Output measurement in the service sectors*". Chicago: University of Chicago Press, NBER, 245-279.
- BERGER, Allen N. & MESTER, Loretta J. (1997), Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of Banking and Finance*, **21**, 895-947.
- CAMPOS, Maurício B. (2002), *Produtividade e eficiência do setor bancário privado brasileiro de 1994 a 1999*. Dissertação de mestrado apresentada ao curso de pós-graduação da EAESP/FGV.
- CHABALGOITY, Luiz, MARINHO, Emerson, BENEGAS, Maurício, JORGE NETO, Paulo de M. (2005). The impact of de-regulamentation in the Brazilian banking industry: a production metafrontier approach. In: ASSOCIAÇÃO NACIONAL DOS CENTROS DE PÓS-GRADUAÇÃO EM ECONOMIA, *Anais do XXXIII Encontro Nacional de Economia*.
- COELLI, Tim J., RAO, D.S. Prasada & BATTESE, George E. (1998). *An Introduction to Efficiency and Productivity Analysis*. Boston: Kluwer Academic Publishers.
- COSTA, Ana Carla A. & NAKANE, Márcio I. (2004) *Revisitando a metodologia de decomposição do spread bancário no Brasil*. Seminários Bacen-USP de Economia Monetária e Bancária – 29/10/2004.

FIXLER, Dennis J. & ZIESCHANG, Kimberly D. (1992) User costs, shadow prices, and the real output of banks. In: GRILICHES, Zvi. (ed.), *Output measurement in the service sectors*. Chicago: University of Chicago Press, NBER, 219-243.

HANCOCK, D. (1985), The financial firm: production with monetary and nonmonetary goods. *Journal of Political Economy*, **93**, 859-880.

HUMPHREY, David B. (1992), Flow versus stock indicators of bank output: effects on productivity and scale economies measurement. *Journal of Financial Services Research*, **6**, 115-135.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (2004), Contas nacionais trimestrais. *Série Relatórios Metodológicos*, **28**, Rio de Janeiro: IBGE.

JORGE NETO, Paulo de M., ARAÚJO, Luiz Alberto d'A. de, PONCE, David A. S. (2005). Competição e concentração entre bancos brasileiros. In: ASSOCIAÇÃO NACIONAL DOS CENTROS DE PÓS-GRADUAÇÃO EM ECONOMIA, *Anais do XXXIII Encontro Nacional de Economia*.

KRAUSE, Kathleen & TABAK, Benjamin M. (2004), Eficiência bancária: uma aplicação DEA para os segmentos bancários no Brasil. In: BANCO CENTRAL DO BRASIL, *Relatório de Estabilidade Financeira – Maio/2004*.

KUMBHAKAR, Subal C. & LOVELL, C. A. Knox (2000), *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

GALLANT, A. Ronald (1981), On the Bias of flexible functional forms and an essentially unbiased form: The Fourier flexible form. *Journal of Econometrics*, **15**, 211-245.

_____ (1982), Unbiased determination of production technologies. *Journal of Econometrics*, **20**, 285-324.

LEVINE, Ross (1997), Financial development and economic growth: views and agenda. *Journal of Economic Literature*, **35**, 688-726.

MITCHELL, Karlyn & ONVURAL, Nur M. (1996), Economies of scale and scope at large commercial banks: Evidence from the Fourier flexible functional form. *Journal of Credit, Money and Banking*, **28**, 178-199.

NAKANE, Márcio I. (1999), *Productive efficiency in the Brazilian banking sector*. Texto para Discussão 20/99, São Paulo: IPE-USP.

_____ (2002), A test of competition for Brazilian banking. *Estudos Econômicos*, **32**, 203-224.

_____ & WEINTRAUB, Daniela B. (2005), Bank privatization and productivity: evidence for Brazil. *Journal of Banking and Finance*, **29**, 2259-2289.

PESEK, Boris P. (1970), Bank's supply function and the equilibrium quantity of money. *Canadian Journal of Economics*, **3**, 357-385.

SILVA, Tarcio L. da & JORGE NETO, Paulo de M. (2002), Economias de escala e eficiência nos bancos brasileiros após o Plano Real. *Estudos Econômicos*, **32**,577-619.

SOUSA, Geraldo da S., STAUB, Roberta B. & TABAK, Benhamin M. (2003), Avaliação da significância do efeito de fatores nas medidas de eficiência da DEA orientada a produto: aplicação para bancos brasileiros. *Relatório de Estabilidade Financeira – Novembro/2003*.

STATACORP (2005). *Stata Base Reference Manual: Volume 1, A-J*. College Station: Stata Press.

TRIPLETT, Jack E. (1992), Comment on “Measurement and efficiency issues in commercial banking”. In: GRILICHES, Zvi. (ed.), *Output measurement in the service sectors*. Chicago: University of Chicago Press, NBER, 287-296.

UNITED NATIONS (1993), *A system of national accounts 1993*. New York: United Nations. [disponível em <http://unstats.un.org/unsd/sna1993/toctop.asp>]

WANG, J. Christina (2003a), *Loanable funds, risk, and bank service output*. Federal Reserve Bank of Boston Working Paper 03-4.

_____ (2003b), *Service output and bank holding companies in the 1990s and the role of risk*. Federal Reserve of Boston Working Paper 03-5.

_____ (2003c), *Productivity and economies of scale in the production of bank service value added*. Federal Reserve Bank of Boston Working Paper 03-6.

_____ (2003d), *Merger-related cost savings in the production of banking services*. Federal Reserve Bank of Boston Working Paper 03-7.

WYKOFF, Frank C. (1992), Comment on “Measurement and efficiency issues in commercial banking”. In: GRILICHES, Zvi. (ed.), *Output measurement in the service sectors*. Chicago: University of Chicago Press, NBER, 279-287.

APÊNDICE 1 – CÁLCULO DAS VARIÁVEIS USADAS NO CAPÍTULO 3

Este apêndice apresenta em maior detalhe a construção das variáveis utilizadas no capítulo 3.

Primeiramente, o produto bancário calculado é dado pela soma do produto explícito (Rendas de Prestação de Serviços) e o produto implícito das divisões 1 e 3, definidos no Quadro A.1 (que, por sua vez, reproduz o Quadro 3.1).

Quadro A.1: Mensuração do Produto Implícito Bancário

Divisão 1 – Crédito	Divisão 3 - Depósitos
Rendas de Operações de Crédito	(Taxa Selic)*(Saldo de Depósitos)
- (Taxa Selic)*(Saldo de Operações de Crédito)	
- (Alíquota FGC)*(Saldo de Operações de Crédito)	- Despesas de Captação
- (Prêmio de Risco)*(Saldo de Operações de Crédito)	
= Produto Implícito da Divisão 1	= Produto Implícito da Divisão 3

A alíquota do FGC aplicada é de 0,025% a.m. O prêmio de risco é dado por:

$$r^{RP} = \frac{Br^F}{(B - I)} - r^F \quad (A1)$$

Onde r^F é a taxa livre de risco (Selic), B é o saldo das operações de crédito e I iguala 40% do saldo das operações de crédito com mais de 90 dias de atraso. Onde há referência a “operações de crédito”, inclui-se também as “operações de arrendamento mercantil”.

Os códigos dos itens contábeis do COSIF utilizados no cálculo são:

- Rendas das operações de crédito: 7.1.1.00.00-1
- Rendas de operações de arrendamento mercantil: 7.1.2.00.00-4
- Saldo das operações de crédito: 1.6.0.00.00-1
- Saldo das operações de arrendamento mercantil: 1.7.0.00.00-0
- Saldo de depósitos iguala a soma de:
 - Depósitos à vista: 4.1.1.00-00-0
 - Depósitos em caderneta de poupança: 4.1.2.00.00-3
 - Depósitos a prazo: 4.1.5.00.00-2

- Despesas de captação igualam à soma de:
 - Despesas com depósitos a prazo: 8.1.1.30.00-9
 - Despesas com depósitos em caderneta de poupança: 8.1.1.10.00-5
 - Despesas com depósitos a prazo com reaplicação automática: 8.1.1.35.00-4
- Rendas de prestação de serviços: 7.1.7.00.00-9
- Saldo de operações de crédito com mais de 90 dias de atraso iguala a soma de:
 - Operações de crédito – classificação nível E: 3.1.6.00.00-8
 - Operações de crédito – classificação nível F: 3.1.7.00.00-1
 - Operações de crédito – classificação nível G: 3.1.8.00.00-4
 - Operações de crédito – classificação nível H: 3.1.9.00.00-7

A variável medindo o insumo *trabalho* é dada pela soma do número de “*funcionários*”, “*membros do conselho*” e “*diretores*”, disponíveis no Sistema de Informações do Banco Central - SISBACEN. Seu preço é dado pela razão entre as *despesas de pessoal* e a variável *trabalho*, onde a primeira é formada pela soma de:

- Despesas de honorários: 8.1.7.18.00-5
- Despesas de pessoal – benefícios: 8.1.7.27.00-3
- Despesas de pessoal – encargos sociais: 8.1.7.30.00-7
- Despesas de pessoal – proventos: 8.1.7.33.00-4
- Despesas de pessoal – treinamento: 8.1.7.36.00-1
- Despesas de remuneração de estagiários: 8.1.7.37.00-0

O insumo *despesas administrativas* é dado pela soma de:

- Despesas de água, energia, e gás: 8.1.7.03.00-3
- Despesas de comunicações: 8.1.7.12.00-1
- Despesas de manutenção e conservação de bens: 8.1.7.21.00-9
- Despesas de material: 8.1.7.24.00-6
- Despesas de processamento de dados: 8.1.7.39.00-8

O preço do insumo *despesas administrativas* é dado pela razão entre este e o *Ativo Circulante e Realizável a Longo Prazo* (1.0.0.00.00-7).³⁷

O insumo *capital físico* é calculado pela soma de:

- Imóveis de uso: 2.2.3.00.00-1
- Instalações, móveis e equipamentos de uso: 2.2.4.00.00-4
- Item “Outros” do imobilizado de uso: 2.2.9.00.00-9
- Imobilizado de arrendamento: 2.3.0.00.00-1

O *custo variável* é formado pela soma entre as despesas com trabalho e as despesas administrativas. Abaixo estão listados os itens COSIF usados para gerar as variáveis determinantes da ineficiência:

- *Ativo* iguala o logaritmo natural do total de ativos, que é dado pela soma de:
 - Ativo circulante e realizável a longo prazo: 1.0.0.00.00-7
 - Ativo permanente: 2.0.0.00.00-4
- *Plativo* iguala a razão entre o Patrimônio Líquido (6.0.0.00.00-2) e o total de ativos.
- *Empativo* é dado pela razão entre a soma do saldo de fundos emprestados e o total de ativos, onde o primeiro é dado por:
 - Obrigações por operações compromissadas: 4.2.0.00.00-6
 - Recursos de aceites cambiais, letras imobiliárias e hipotecárias, debêntures e similares: 4.3.0.00.00-5
 - Obrigações por empréstimos e repasses: 4.6.0.00.00-2
- *Depav* é dado pela razão entre o saldo de depósitos à vista (4.1.1.00.00-0) e o total de ativos.
- *Nonperform* é dado pela razão entre o saldo de operações de crédito com mais de 90 dias de atraso e o saldo total de operações de crédito, sendo que ambos estão definidos acima.
- As variáveis *dummy* de estrutura de propriedade (*cestrang*, *pestrang* e *publico*) foram calculadas com base em informações disponíveis no *site* do Banco Central do Brasil.

Por fim, o Quadro A.2 lista os bancos incluídos na amostra.

³⁷ A fim de reportar as estatísticas descritivas desta variável no Quadro 3.1 sem o uso de casas decimais, foi utilizada esta razão multiplicada por 10⁶.

Quadro A.2: Bancos da Amostra

BCO. DE LA NACION ARGENTINA	BCO. MODAL
BCO. A. J. RENNER	BCO. MORADA S.A
BCO. ABN AMRO REAL	BCO. OURINVEST
BCO. ALFA	BCO. PACTUAL
BCO. BARCLAYS	BCO. PANAMERICANO
BCO. BBM S/A	BCO. PECUNIA
BCO. BEC	BCO. PSA FINANCE BRASIL
BCO. BGN	BCO. RENDIMENTO
BCO. BMG	BCO. SAFRA
BCO. BNP PARIBAS BRASIL	BCO. SANTANDER BRASIL
BCO. BONSUCESSO	BCO. SCHAHIN
BCO. BRADESCO	BCO. SOCIETE GENERALE BRASIL
BCO. CACIQUE	BCO. SOFISA
BCO. CALYON BRASIL	BCO. SUDAMERIS BRASIL
BCO. CAPITAL	BCO. TRIANGULO
BCO. CITIBANK	BCO. VR
BCO. CRUZEIRO DO SUL	BCO. COOPERATIVO DO BRASIL
BCO. DE PERNAMBUCO - BANDEPE	BCO. DO ESTADO DO ESPIRITO SANTO
BCO. DE TOKYO-MITSUBISHI	BANIF - BCO. INTERNACIONAL DO FUNCHAL
BCO. DIBENS	BANKBOSTON BCO. MULTIPLO
BCO. DO BRASIL	BCO. COOPERATIVO SICREDI
BCO. DO ESTADO DE SANTA CATARINA	BASA BCO. DA AMAZONIA
BCO. DO ESTADO DE SÃO PAULO	BCO. COMERCIAL URUGUAI
BCO. DO ESTADO DE SERGIPE	BNB - BCO. DO NORDESTE
BCO. DO ESTADO DO PARÁ	BRB - BCO. DE BRASILIA
BCO. DO ESTADO DO PIAUÍ - BEP	CAIXA ECONÔMICA FEDERAL
BCO. DO ESTADO DO RIO GRANDE DO SUL	DEUTSCHE BANK BCO. ALEMAO
BCO. FIAT	BCO. GUANABARA
BCO. FICSA	HSBC BANK BRASIL - BCO. MULTIPLO
BCO. GE CAPITAL	BCO. KEB DO BRASIL
BCO. GERDAU S.A	BCO. DE LA REP. ORIENTAL DEL URUGUAY
BCO. GMAC S/A	LEMON BANK BCO. MULTIPLO
BCO. INDUSTRIAL DO BRASIL	BCO. LUSO BRASILEIRO
BCO. INDUSTRIAL E COMERCIAL	NOSSA CAIXA
BCO. INTERCAP	PARANA BANCO
BCO. INVESTCRED UNIBCO.	BCO. RIBEIRÃO PRETO
BCO. JOHN DEERE	BCO. SUMITOMO MITSUI
BCO. MÁXIMA	UNIBANCO
BCO. MERCANTIL DO BRASIL	

APÊNDICE 2 – RESULTADOS AUXILIARES DA ESTIMAÇÃO DAS FRONTEIRAS DE PRODUÇÃO E CUSTO

Este apêndice apresenta os resultados dos modelos “irrestritos”, i.e., sem a exclusão de nenhuma variável para as fronteiras de produção e custo. O Quadro A.3 apresenta os resultados para a fronteira de produção.

Quadro A.3: Resultados da Fronteira de Produção – Modelo Irrestrito

Variável	Coefficiente	Erro-Padrão	P-Valor	Variável	Coefficiente	Erro-Padrão	P-Valor
$\ln(X_1)$	12.009	1.774	0.000	$\sin(2Z_1)$	-0.841	0.224	0.000
$\ln(X_2)$	-0.116	2.649	0.965	$\sin(2Z_2)$	-0.325	0.211	0.123
$\ln(X_3)$	-11.743	4.987	0.019	$\sin(2Z_3)$	-2.063	0.413	0.000
$\ln(X_1)^2$	-0.753	0.248	0.002	$\sin(Z_1+Z_2)$	-0.061	0.270	0.821
$\ln(X_2)^2$	0.037	0.157	0.815	$\sin(Z_2+Z_3)$	1.436	0.449	0.001
$\ln(X_3)^2$	0.992	0.301	0.001	$\sin(Z_1+Z_3)$	1.951	0.461	0.000
$\ln(X_1)\ln(X_2)$	-0.074	0.037	0.045	$\ln(X_1)T$	-0.009	0.005	0.097
$\ln(X_1)\ln(X_3)$	-0.303	0.071	0.000	$\ln(X_2)T$	0.005	0.005	0.330
$\ln(X_2)\ln(X_3)$	0.020	0.070	0.773	$\ln(X_3)T$	0.003	0.006	0.575
$\cos(Z_1)$	7.333	1.645	0.000	λ_2	0.017	0.101	0.868
$\cos(Z_2)$	0.152	1.538	0.921	λ_3	-0.125	0.150	0.406
$\cos(Z_3)$	-8.685	3.499	0.013	λ_4	-0.436	0.208	0.037
$\cos(2Z_1)$	1.825	0.240	0.000	λ_5	-0.370	0.268	0.167
$\cos(2Z_2)$	-0.123	0.345	0.722	λ_6	-0.404	0.330	0.221
$\cos(2Z_3)$	0.075	0.681	0.912	λ_7	-0.590	0.393	0.133
$\cos(Z_1+Z_2)$	-0.177	0.314	0.573	λ_8	-0.601	0.456	0.187
$\cos(Z_1+Z_3)$	-2.708	0.437	0.000	λ_9	-0.659	0.519	0.204
$\cos(Z_2+Z_3)$	0.138	0.471	0.770	λ_{10}	-0.738	0.582	0.205
$\sin(Z_1)$	3.738	0.436	0.000	λ_{11}	-0.772	0.647	0.233
$\sin(Z_2)$	1.211	0.516	0.019	λ_{12}	-0.525	0.711	0.461
$\sin(Z_3)$	-1.671	1.059	0.115	constante	48.005	40.877	0.240

Oito variáveis do modelo irrestrito para a fronteira de produção foram retiradas. A estatística de Wald para o teste de significância conjunta destas variáveis foi 3.78, de tal modo que não é possível rejeitar a hipótese nula de que o modelo mais adequado é o reportado no Quadro 3.3.

Quadro A.4: Resultados da Fronteira de Produção – Modelo Irrestrito

Variável	Erro-			Variável	Erro-		
	Coefficiente	Padrão	P-Valor		Coefficiente	Padrão	P-Valor
$\ln(W_1)$	2.064	2.512	0.411	$\text{sen}(Z_2)$	-0.635	0.181	0.000
$\ln(K)$	5.545	1.039	0.000	$\text{sen}(Z_3)$	0.382	0.149	0.010
$\ln(Y)$	-4.260	1.257	0.001	$\text{sen}(Z_4)$	-1.559	0.225	0.000
$\ln(W_2)$	-3.966	2.837	0.162	$\text{sen}(2Z_1)$	-0.220	0.092	0.017
$\ln(W_1)^2$	-0.183	0.172	0.288	$\text{sen}(2Z_2)$	-0.227	0.052	0.000
$\ln(K)^2$	-0.267	0.067	0.000	$\text{sen}(2Z_3)$	-0.476	0.063	0.000
$\ln(Y)^2$	0.417	0.078	0.000	$\text{sen}(2Z_4)$	-0.818	0.092	0.000
$\ln(W_2)^2$	0.801	0.459	0.081	$\text{sen}(Z_1+Z_2)$	0.170	0.018	0.000
$\ln(W_1)\ln(K)$	-0.006	0.008	0.451	$\text{sen}(Z_2+Z_4)$	0.021	0.014	0.145
$\ln(K)\ln(Y)$	-0.067	0.019	0.000	$\text{sen}(Z_2+Z_3)$	-0.005	0.012	0.667
$\ln(W_1)\ln(Y)$	0.020	0.017	0.237	$\text{sen}(Z_1+Z_3)$	-0.047	0.013	0.000
$\ln(W_1)\ln(K)$	-0.034	0.013	0.008	$\text{sen}(Z_1+Z_4)$	0.049	0.021	0.018
$\ln(W_2)\ln(W_2)$	0.064	0.043	0.137	$\text{sen}(Z_3+Z_4)$	0.805	0.119	0.000
$\ln(W_2)\ln(Y)$	-0.008	0.050	0.867	$\ln(W_1)T$	0.003	0.003	0.308
$\cos(Z_1)$	2.164	1.427	0.129	$\ln(W_2)T$	0.008	0.004	0.048
$\cos(Z_2)$	-1.450	0.885	0.101	$\ln(K)T$	0.000	0.002	0.927
$\cos(Z_3)$	3.252	0.622	0.000	$\ln(Y)T$	-0.001	0.002	0.568
$\cos(Z_4)$	-2.975	0.766	0.000	λ_2	-0.033	0.070	0.635
$\cos(2Z_1)$	0.426	0.204	0.037	λ_3	-0.104	0.130	0.423
$\cos(2Z_2)$	-0.080	0.120	0.507	λ_4	-0.176	0.193	0.362
$\cos(2Z_3)$	0.613	0.099	0.000	λ_5	-0.246	0.257	0.338
$\cos(2Z_4)$	-0.487	0.133	0.000	λ_6	-0.326	0.321	0.311
$\cos(Z_1+Z_2)$	-0.014	0.017	0.418	λ_7	-0.398	0.385	0.302
$\cos(Z_2+Z_4)$	-0.021	0.014	0.119	λ_8	-0.447	0.450	0.320
$\cos(Z_2+Z_3)$	-0.010	0.014	0.454	λ_9	-0.502	0.513	0.328
$\cos(Z_1+Z_3)$	0.024	0.017	0.152	λ_{10}	-0.565	0.578	0.328
$\cos(Z_1+Z_4)$	-0.068	0.016	0.000	λ_{11}	-0.593	0.644	0.358
$\cos(Z_3+Z_4)$	-0.116	0.125	0.355	λ_{12}	-0.727	0.709	0.305
$\text{sen}(Z_1)$	-0.804	0.354	0.023	constante	-11.029	17.353	0.525

Sete variáveis do modelo irrestrito para a fronteira de custo foram retiradas. A estatística de Wald para o teste de significância conjunta destas variáveis foi 3.93, de tal modo que não é possível rejeitar a hipótese nula de que o modelo mais adequado é o reportado no Quadro 3.6.