



BANCO CENTRAL DO BRASIL

Trabalhos para Discussão

13

Modelos de Previsão de Insolvência Bancária no Brasil

Marcio Magalhães Janot

Março/2001

ISSN 1519-1028

CGC 00.038.166/0001-05

Trabalhos para Discussão	Brasília	nº 13	Mar	2001	P. 1 – 41
--------------------------	----------	-------	-----	------	-----------

Trabalhos para Discussão

Editado por:

Departamento de Estudos e Pesquisas (Depep)

(e-mail: conep.depep@bcb.gov.br , workingpaper@bcb.gov.br)

Reprodução permitida somente se a fonte for citada como: Trabalhos para Discussão nº 13.

Autorizado por Ilan Goldfajn (Diretor de Política Econômica).

Controle Geral de Assinaturas:

Banco Central do Brasil
Demap/Disud/Subip
SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo
70074-900 - Brasília (DF)
Telefone (61) 414-1392
Fax (61) 414-3165

Tiragem: 450 exemplares

The views expressed in this work are those of the authors and do not reflect those of the Banco Central or its members.

Although these Working Papers often represent preliminary work, citation of source is required when used or reproduced.

As opiniões expressas neste trabalho são exclusivamente do(s) autor(es) e não refletem a visão do Banco Central do Brasil.

Ainda que este artigo represente trabalho preliminar, citação da fonte é requerida mesmo quando reproduzido parcialmente.

Central de Informações do Banco Central do Brasil

Endereço: Secre/Surel/Dinfo
Edifício-Sede, 2º subsolo
SBS - Quadra 3, Zona Central
70074-900 - Brasília (DF)
Telefones: (61) 414 (...) 2401, 2402, 2403, 2404, 2405, 2406
DDG: 0800 992345
Fax: (61) 321 9453
Internet: <http://www.bcb.gov.br>
E-mail: cap.secre@bcb.gov.br
dinfo.secre@bcb.gov.br

Modelos de Previsão de Insolvência Bancária no Brasil

Marcio Magalhães Janot*

Resumo

Este estudo examina a eficácia de dois tipos de modelos de *early warning* - o modelo de regressão logística e o modelo de risco proporcional de Cox - em prever o fenômeno de insolvência bancária no Brasil durante o período 1995/1998. Estes modelos basicamente produzem estimativas da probabilidade de um banco, com um dado conjunto de características, sobreviver mais que um determinado intervalo de tempo no futuro, classificando-o como solvente ou insolvente. Apontam também quais as características que mais contribuíram para a insolvência das instituições financeiras. O alto percentual de acerto de classificação dos bancos pelos dois modelos estimados, com a identificação de uma proporção considerável das insolvências com antecedência, indicam que a insolvência bancária é passível de ser prevista no Brasil, sendo recomendável a utilização destes como um instrumento adicional de supervisão do sistema financeiro pelo Banco Central.

* Central Bank of Brazil.

Modelos de Previsão de Insolvência Bancária no Brasil[◇]

1. Introdução

O sistema financeiro brasileiro vem passando por uma profunda transformação desde a implantação do Real como a nova moeda nacional em 1º de julho de 1994. Com a redução significativa da inflação, os bancos não tinham mais como contar com as receitas provenientes do *floating*. Para sobreviver, os bancos passaram a buscar fontes alternativas de receitas. O acirramento da concorrência, inclusive com a entrada de grandes bancos estrangeiros no país, fizeram com que muitos bancos tivessem que passar por algum tipo de ajuste como transferência de controle acionário, intervenção e/ou liquidação por parte do Banco Central.

Com essa mudança de ambiente macroeconômico cresceu a preocupação dos agentes econômicos em mensurar a solidez das instituições financeiras individualmente e em conjunto. Grandes depositantes e outros credores estão interessados em mensurar o risco de perda do capital aplicado, enquanto o agente regulador e supervisor em antecipar problemas, evitando crises sistêmicas e grandes prejuízos aos credores e ao país. Embora ainda exista alguma controvérsia sobre quais os fatores que predominam na causalidade das crises bancárias, há um consenso na literatura de que ambos os desequilíbrios macroeconômicos e as deficiências a nível microeconômico, como gerenciamento bancário deficiente e falhas na legislação estão no centro das crises bancárias.

A evidência empírica de crises bancárias nos países desenvolvidos e emergentes tem mostrado que nem todos os bancos se tornaram insolventes ou ao menos passaram por dificuldades. Realmente, a experiência latino americana conta muitos casos de bancos grandes que ficaram ainda mais fortes após a crise sistêmica. Portanto, embora os

[◇] Agradeço aos professores Márcio Garcia, Cristiano Fernandes e Dionísio Carneiro Dias pelos comentários e sugestões, indispensáveis para a realização deste trabalho. Todas as conclusões são de responsabilidade do autor, não refletindo necessariamente a visão do Banco Central do Brasil.

fatores macroeconômicos permaneçam importantes em antecipar crises bancárias sistêmicas, é também crucial identificar quais bancos ou grupo de bancos são mais propensos a crise.

Nesse contexto, é de suma importância para os órgãos governamentais responsáveis pela formulação de políticas e supervisão do sistema financeiro que se crie um “sistema de alerta” baseado em variáveis microeconômicas que deixe o governo em uma melhor posição para prever a ocorrência de uma crise bancária ao menor custo possível, combinando políticas macroeconômicas com outras designadas a afetar somente aquelas instituições julgadas mais frágeis e propensas a se tornarem insolventes. Para enfatizar esse ponto, considere uma situação onde o governo não tenha tal sistema e tome decisões baseado somente em indicadores macroeconômicos. Assuma que estes indicadores estejam sinalizando uma alta probabilidade de uma crise num futuro próximo. Sem saber quais instituições estão mais frágeis, as opções de política ficam limitadas aquelas que corrijam problemas a nível agregado, afetando tanto os bancos fracos como os saudáveis. Isso talvez imponha restrições desnecessárias nas fontes de financiamento da atividade econômica.

Esse trabalho apresenta dois “sistemas de alerta” construídos a partir da aplicação de técnicas estatísticas para analisar os demonstrativos financeiros dos bancos. Esses sistemas baseados em variáveis microeconômicas tem por objetivo mensurar o risco de insolvência de cada banco individualmente. O objetivo é construir um modelo de previsão de insolvência bancária - chamado na literatura internacional de *early warning model* - que expresse a probabilidade do banco se tornar insolvente no futuro como uma função das variáveis obtidas do balanço patrimonial e do demonstrativo de resultado do corrente período.

O primeiro passo a ser dado na construção do modelo é definir o que constitui uma insolvência bancária. Como qualquer outra empresa, um banco se torna insolvente quando seu patrimônio líquido se torna negativo ou se for impossível continuar suas operações sem incorrer em perdas que resultariam em patrimônio líquido negativo. No entanto, muitos casos de insolvência bancária são resolvidos através de fusões supervisionadas

entre o banco insolvente e uma instituição saudável, não havendo uma falência no sentido legal. Assim, podemos definir insolvência de diferentes maneiras. O critério adotado na maioria dos estudos, inclusive o presente, é classificar como insolventes os bancos que sofreram intervenção ou liquidação extrajudicial por parte do Banco Central. Neste caso, identificamos este grupo de bancos e suas características um ou mais anos anteriores à liquidação e comparamos com um grupo de bancos solventes. Outra forma de definir insolvência seria quando o banco é colocado em evidência pelo supervisor responsável. Este tipo de abordagem apresenta a desvantagem de depender de critérios subjetivos de julgamento. Por outro lado, como a colocação de um banco em evidência precede sua liquidação, um modelo que preveja esta classificação provê um tempo maior para ações corretivas.

A principal contribuição deste trabalho para a literatura de insolvência bancária no Brasil é a metodologia. O critério de seleção dos indicadores econômico-financeiros foi a aplicação do teste t sobre cada indicador, sendo selecionados os indicadores significantes ao nível de 5%. Além disso, é testada a validade dos modelos tanto para os bancos que compõe as amostras quanto para os demais bancos do sistema financeiro, de forma segmentada por origem do capital e por porte. A estabilidade dos modelos estimados também é testada ao utilizarmos, como variáveis independentes, indicadores econômico-financeiros coletados de anos subsequentes ao ano em que o modelo foi estimado.

2. Fundamentação Teórica

2.1 O Modelo Logit

Seja N o tamanho da amostra de bancos a serem analisados; M o número de indicadores econômico-financeiros selecionados como variáveis explicativas; i indexa os bancos, $i \in (1, N)$; j indexa as variáveis, $j \in (1, M)$; x_{ij} o valor da variável j para o banco i ; e $Y = Y_1, \dots, Y_n$ variável dependente dicotômica: $Y_i = 1$ para os bancos que faliram e $Y_i = 0$ para os não falidos.

Como o objetivo do modelo é detectar problemas em instituições financeiras com antecedência, possibilitando a instituição supervisora a tomada de medidas corretivas que evitem a falência ou pelo menos reduzam o seu custo, as variáveis explicativas são construídas a partir de uma base de dados de um período anterior àquele da variável dependente Y .

O modelo empírico de previsão de insolvência bancária *ex-post* pode ser escrito na sua forma mais genérica como

$$\Pr(Y_i = 1) = F(x_{i1}, x_{i2}, \dots, x_{iM}, b_1, b_2, \dots, b_m). \quad (2.1)$$

ou seja, a probabilidade de insolvência do banco i ($Y_i = 1$) é função dos indicadores x_{ij} 's coletados das demonstrações financeiras de uma data anterior a efetiva insolvência e dos coeficientes b_1, b_2, \dots, b_m que são constantes estimadas a partir dos dados amostrais. A forma da função de distribuição de probabilidade F não é um resultado empírico e sim uma suposição, embora especificações alternativas de F deveriam ser testadas nos dados para a seleção da forma mais apropriada.

Uma variedade de técnicas estatísticas multivariadas podem ser utilizadas para prever uma variável dependente dicotômica a partir de um conjunto de variáveis independentes como, por exemplo, a análise de regressão múltipla e a análise discriminante. Quando a variável dependente pode assumir apenas dois valores, as suposições necessárias para testar hipóteses na análise de regressão múltipla são violadas. Por exemplo, não é razoável assumir que a distribuição dos erros seja normal. Outra dificuldade na análise da regressão múltipla é que os valores previstos não podem diretamente ser interpretados como probabilidades por não estarem restritos a ficar dentro do intervalo entre 0 e 1.

A análise discriminante linear permite uma previsão direta do grupo que a variável pertence. No nosso trabalho, isso corresponde a classificar os bancos como pertencendo ao grupo de bancos solventes ou ao grupo de bancos insolventes. No entanto, esta

técnica requer a suposição de normalidade multivariada das variáveis independentes e matrizes de variância-covariância iguais nos dois grupos para que a regra de previsão seja ótima.

O modelo de regressão logística requer bem menos suposições que a análise discriminante e, mesmo quando as suposições requeridas para a análise discriminante são satisfeitas, a regressão logística continua apresentando bons resultados¹. O modelo logit é um modelo de resposta qualitativa, pois é utilizado com o propósito de modelar o comportamento de um tomador de decisão que deve escolher entre um conjunto finito de alternativas. Estes modelos são aplicáveis a um conjunto mais extenso de situações de pesquisa que a análise discriminante². Exemplos de situações na qual estes modelos tem sido usados incluem a decisão de um indivíduo entre comprar ou alugar uma casa, a decisão de um senador se vota sim ou não na votação de uma determinada lei, a decisão da instituição supervisora se liquida ou não um banco.

Embora a análise discriminante e os modelos de resposta qualitativa possam ser utilizados de modo trocável, as motivações no uso dos dois modelos são bem diferentes. O modelo de análise discriminante especifica uma distribuição conjunta da variável dependente (Y_i) e das variáveis independentes (x_i), não apenas a distribuição condicional de Y_i dado x_i . Nos modelos de resposta qualitativa, a determinação de x_i (características dos bancos) claramente precede a de Y_i (insolvência); dessa forma, é importante especificar $\Pr(Y_i=1 / X)$, enquanto a especificação da distribuição de X pode ser ignorada. Ao contrário, no modelo de análise discriminante, a afirmativa Y_i precede a determinação de X . Em síntese, a análise discriminante é meramente uma técnica de classificação, enquanto o modelo logit analisa uma relação causal. Como não estamos interessados apenas numa classificação dicotômica dos bancos, mas também na estimação dos coeficientes das variáveis independentes e, conseqüentemente, na determinação da probabilidade de ocorrência de insolvência bancária, o modelo logit é o mais apropriado ao nosso estudo.

¹ Norius (1993).

² Judge et al.(1985).

A forma funcional assumida da equação (2.1) no modelo logit é a função logística dada por

$$\Pr(Y_i = 1) = P_i = \frac{1}{1 + e^{-W_i}}, i = 1, \dots, N \quad (2.2)$$

onde $W_i = b_0 + \sum_{j=1}^M b_j x_{ij}$ é uma combinação linear das variáveis independentes e do conjunto de coeficientes $B = (b_0, b_1, \dots, b_M)$ que serão estimados.

Esta especificação pode ser derivada de um modelo de insolvência bancária, onde as variáveis independentes são indicadores econômico-financeiros ou outros dados supostamente relevantes para determinar o risco de insolvência de uma instituição financeira. A relação entre os indicadores e a probabilidade estimada é não linear. A probabilidade estimada vai estar sempre entre 0 e 1, independentemente do valor de W. Assume-se que exista uma combinação linear W destas variáveis independentes que é positivamente relacionada com a probabilidade de insolvência. Ou seja, quanto maior o valor de W_i , maior a probabilidade de insolvência do banco i, condicional aos valores dos indicadores do banco. Assim, podemos dizer que W é um índice da propensão a falir de um banco. No entanto, não é realista esperar que qualquer conjunto selecionado de variáveis independentes possa prover com exatidão a condição econômico-financeira dos bancos e que possamos prever com certeza se a instituição financeira vai falir ou não. Existem incontáveis influências no resultado que somente são observadas *ex-post* como a qualidade da carteira de empréstimos e eventos puramente aleatórios que determinam, em termos das variáveis incluídas, qual a propensão a falir que um banco deve ter para realmente falir. Ou seja, cada banco tem uma “tolerância para vulnerabilidade” \bar{W}_i de tal modo que se $W_i \geq \bar{W}_i$ então o banco irá falir, caso contrário não. \bar{W}_i é uma função das variáveis excluídas, e por definição é não observável. Nós só

sabemos *ex-post* se o valor de W_i de cada banco era maior ou menor que o respectivo \bar{W}_i ³.

Supondo que \bar{W} é uma variável aleatória com função de distribuição de probabilidade $G(w)$, podemos escrever a probabilidade de insolvência *ex-ante* como uma função de W_i da seguinte forma:

$$\Pr(Y_i = 1) = \Pr(W_i \geq \bar{W}_i) = G(W_i) \quad (2.3)$$

A hipótese feita acima é justificável, já que \bar{W} é uma combinação de varias variáveis sendo muitas delas desconhecidas e aleatórias. Usando a distribuição normal padrão na especificação de G para estimar os coeficientes de W_i estaríamos utilizando um modelo conhecido como *probit*. Em muitas aplicações práticas torna-se desejável o uso de uma aproximação da distribuição normal ao invés da normal propriamente dita. A distribuição logística descrita na equação (2.2) fornece esta aproximação, distinguindo-se da normal somente nos extremos da cauda⁴.

Para melhor interpretação dos coeficientes logísticos, reescrevemos a regressão logística em termos de *odds* de um evento. O *odds* de um evento é definido como o quociente entre a probabilidade que o evento ocorra e seu complementar. Por exemplo, o *odds* de obter uma cara no lançamento de uma moeda é $0,5 / 0,5 = 1$. O modelo logístico em termos do log de *odds* é chamado de logit e pode ser escrito como:

$$\log\left(\frac{P_i}{1 - P_i}\right) = b_0 + b_1 x_{i1} + \dots + b_M x_{iM} = W_i. \quad (2.4)$$

Pela equação (2.4), nota-se que os coeficientes da regressão logística podem ser interpretados como a mudança no log de *odds* associado com uma mudança unitária na

³ Neste capítulo, chamamos de falência o evento regulamentar de intervenção ou liquidação extrajudicial da instituição financeira pelo Banco Central.

⁴ Martin (1977).

variável independente. Como é mais fácil a interpretação de *odds* do que de log de *odds*, a equação logística pode ser escrita em termos de *odds* como:

$$\frac{P_i}{1 - P_i} = e^{b_o + b_1 x_{i1} + \dots + b_M x_{iM}} = e^{W_i}, \quad (2.5)$$

onde e elevado a potência b_i é o fator no qual o *odds* muda quando a i -ésima variável independente aumenta em uma unidade. Se b_i é positivo, esse fator vai ser maior que 1, o que significa que o *odds* cresceu; se b_i é negativo, o fator vai ser inferior a 1, o que significa que o *odds* decresceu. Quando b_i é 0, o fator iguala 1, deixando o *odds* inalterado. Ou seja, quando o indicador econômico-financeiro apresenta um coeficiente negativo (positivo), significa que qualquer acréscimo no indicador vai representar um aumento (decréscimo) relativo na probabilidade de insolvência.

Essencialmente o que estamos querendo obter ao estimar os coeficientes do modelo logit é produzir um conjunto de estimativas de probabilidade, tendo sido assinalada alta probabilidade *ex-ante* de insolvência aos bancos que faliram, e aqueles que não faliram baixa probabilidade. Um “bom ajuste” é um conjunto de coeficientes que mais se aproxima deste objetivo.

A técnica utilizada para estimar a equação (2.2) é o método de estimação de máxima verossimilhança. Dado o vetor $Y = Y_1, \dots, Y_n$ de resultados, e o vetor de coeficientes $B = b_o, b_1, \dots, b_M$, a função de máxima verossimilhança da amostra de N bancos é dada por:

$$L(Y, B) = \prod_{i=1}^N P_i^{Y_i} (1 - P_i)^{1 - Y_i} \quad (2.6)$$

onde $L(Y, B)$ é escrito como uma função dos coeficientes B e dos resultados atuais Y , já que as probabilidades P_i são determinadas pelos coeficientes. O processo de estimação consiste em encontrar o conjunto de coeficientes que maximize $L(Y, B)$.

A função de máxima verossimilhança, em muitos casos, é altamente não linear nos parâmetros, o que exige a aplicação de algum método de iteração para a obtenção de uma solução. O método de iteração mais comum é o método de Newton-Raphson, que pode ser utilizado para maximizar ou minimizar uma função genérica, não apenas a função de máxima verossimilhança, baseando-se em uma aproximação quadrática da função que desejamos maximizar ou minimizar. Em cada iteração, um conjunto inicial de coeficientes é revisado até o processo convergir para a solução de máxima verossimilhança⁵.

2.2 O Modelo de Risco Proporcional de Cox

A principal vantagem deste modelo em relação aos outros métodos estatísticos utilizados em modelos de previsão de insolvência bancária é que este modelo não produz somente as estimativas da probabilidade de falência ou, alternativamente, de sobrevivência em um determinado período, mas também prevê o tempo esperado para a quebra. Em outras palavras, os modelos de previsão de insolvência baseados em técnicas logit/probit ou análise discriminante nos dá somente a probabilidade estimada de um banco falir em algum ponto no tempo dentro de um intervalo especificado previamente, não provendo nenhuma informação sobre quando a falência vai ocorrer ao longo deste período. Já o modelo de risco proporcional pode ser utilizado para traçar o perfil de sobrevivência para qualquer banco incluído na amostra, ao gerar a probabilidade de sobrevivência estimada para períodos maiores que o especificado como função do tempo. Adicionalmente, este modelo não requer qualquer hipótese sobre as propriedades distributivas dos dados, que talvez pudessem ser violadas. Por exemplo, ao contrário da análise discriminante múltipla, o modelo de risco proporcional não assume a hipótese de normalidade multivariada.

⁵ Para uma descrição mais detalhada do método Newton-Raphson de iteração e das propriedades dos estimadores de máxima verossimilhança, ver Amemiya (1994).

2.2.1 O modelo

O modelo pode ser descrito da seguinte forma. Seja T o tempo até a falência de um determinado banco. Podemos então definir a função de sobrevivência (*survivor function*) dada por $S(t)$ como a probabilidade de um banco sobreviver mais do que t períodos:

$$S(t) = \text{Prob}(T > t) = 1 - F(t). \quad (2.7)$$

onde $F(t)$ é a função de distribuição cumulativa para a variável aleatória tempo de falência. A função de densidade de probabilidade de t é $f(t) = -S'(t)$. Dadas estas definições, podemos especificar a probabilidade de falência no próximo instante, dado que o banco não faliu no período t , como a função risco (*hazard function*):

$$h(t) = \lim_{dt \rightarrow 0} \frac{P(t < T < t + dt / T > t)}{dt} = \frac{-S'(t)}{S(t)}. \quad (2.8)$$

A função de sobrevivência, a função de densidade de probabilidade e a função risco são matematicamente equivalentes, isto é, se uma função é dada, as outras duas podem ser derivadas. No entanto, a distribuição do tempo de falência é mais frequentemente caracterizada pela função risco. Assim, a partir das estimativas de $h(t)$ podemos obter prontamente as estimativas de $S(t)$ através de:

$$S(t) = \exp\left[-\int_0^t h(u)du\right]. \quad (2.9)$$

Diferentes tipos de funções risco podem ser especificadas, dependendo das hipóteses feitas sobre a distribuição do tempo de falência⁶. No modelo de risco proporcional, assume-se que a função risco no tempo t para um banco com um vetor de variáveis explicativas X seja dado por $h(t / X, B) = h_0(t)g(X, B)$, onde os coeficientes B são

⁶ Ver Kiefer (1988) para uma resenha de diferentes tipos de funções de risco.

estimados e descrevem como cada variável explicativa afeta a probabilidade de insolvência. A função $g(\cdot)$ é alguma função de X tal que $g(0) = 1$ e $h_0(t)$ é a função risco de um banco com $X = 0$. Se centralizarmos as variáveis explicativas de tal modo que um banco com $X = 0$ tenha valores iguais às médias populacionais, podemos pensar em $h_0(t)$ como uma função risco para um banco “médio” na população. Assim sendo, a principal suposição na qual se baseia o modelo de risco proporcional é que o efeito das variáveis independentes é multiplicar a função de um banco médio, $h_0(t)$, por alguma função $g(\cdot)$ dos desvios das variáveis explicativas de seus valores médios⁷.

O caso especial dos modelos de risco proporcional que foi utilizado nos trabalhos anteriores sobre o tema é o modelo proposto por Cox (1972). Aplicado extensivamente em estudos na área biomédica, este modelo é particularmente interessante no contexto de prever insolvências bancárias por causa de sua habilidade em prover estimativas do tempo provável de falência e também pela ausência de qualquer hipótese a respeito dos dados. No modelo de Cox, que vai ser usado neste trabalho, temos a função $g(X,B) = \exp(X'B)$. Ou seja, a função risco é dada por:

$$h(t / X, B) = h_0(t) \exp(X' B). \quad (2.10)$$

A principal razão para a escolha de uma função exponencial para $g(\cdot)$ é que tal escolha simplifica imensamente a estimação dos coeficientes B da regressão⁸.

O modelo de risco proporcional de Cox não pode verdadeiramente ser chamado de não paramétrico já que ele depende do vetor de parâmetros da regressão B . Entretanto, a primeira parte da expressão, $h_0(t)$, é o termo não paramétrico denominado função de risco “*baseline*”. Assume-se que $h_0(t)$ é arbitrária e dependente somente do tempo, não sendo necessárias suposições distributivas para realizar a estimação de B ou de $h(t)$.

⁷ O teste desta suposição de proporcionalidade da função risco confirma que a hipótese do risco proporcional é adequada. O teste está descrito em Janot (1999).

⁸ Cox and Oakes (1984).

No presente trabalho o vetor X corresponde a um conjunto de indicadores econômico-financeiros e t corresponde ao tempo de sobrevivência de um banco. A função de sobrevivência a ser estimada aqui é

$$S(t / X, B) = S_0(t)^p \quad , \text{ onde } p = \exp(X'B) \quad (2.11)$$

$$\text{e } S_0(t) = \exp\left[-\int_0^t h_0(u)du\right]$$

é a função de sobrevivência *baseline* correspondente à função de risco *baseline* $h_0(t)$.

Como depende somente do tempo, $S_0(t)$ é igual para todos os bancos. Para calcular a probabilidade de sobrevivência de qualquer banco, é necessário escolher o horizonte temporal que determine a probabilidade *baseline* relevante e então substitui-se os valores das variáveis explicativas na equação (2.11) descrita acima. A função de sobrevivência encontrada nos dará a probabilidade de um banco com certas características sobreviver por t períodos no futuro.

Um ponto a ser questionado no modelo de risco proporcional de Cox é a hipótese de que os valores das variáveis explicativas permanecem constante ao longo do horizonte de tempo especificado na análise. Como, geralmente, os indicadores econômico-financeiros mudam a cada balancete enviado para a instituição reguladora, esta hipótese pode gerar problemas de precisão na classificação.

2.2.2 Amostragem

Em países onde o número de bancos existentes é muito elevado como nos Estados Unidos, a utilização da população inteira de bancos para a geração de um modelo de previsão de insolvência bancária é muito custoso e requer substancial tempo computacional. Portanto, adota-se uma amostra de bancos, que geram modelos de comparável precisão e de muito mais fácil manutenção. Em estudos de insolvência

bancária, a amostragem é uma importante etapa, já que pode significativamente alterar os resultados estimados.

Após a definição da amostra de bancos falidos e da identificação dos bancos que não quebraram para servir de base de comparação e de posse dos indicadores econômico-financeiros que serão testados como possíveis previsores de insolvência, podemos desenvolver o modelo estatístico.

É comum ao estudar a distribuição do tempo de falência, que a amostra consista não somente dos elementos onde o tempo até a falência é conhecido, mas também dos itens para os quais somente um limite inferior para o tempo até a falência seja observado, como é o caso da parte da amostra dos bancos que não faliram. A solução para estes casos é censurar o tempo de sobrevivência destes elementos. Por exemplo, se o período amostral analisado for de 24 meses, o tempo de sobrevivência dos bancos que não quebraram é censurado em 24 meses, já que tudo o que podemos afirmar é que eles sobreviveram no mínimo este tempo no futuro. A possibilidade de censurar tempo de falência é outra vantagem do modelo de risco proporcional.

3. Metodologia

O desenvolvimento de um modelo estatístico de insolvência bancária requer uma seleção prévia de instituições financeiras solventes e insolventes e dos dados contábeis dos bancos selecionados para a construção de indicadores que identifiquem suas características econômico-financeiras.

3.1 Seleção de Instituições Financeiras

A amostra de Instituições Financeiras é composta por um grupo de bancos insolventes, composto pelos 21 bancos que sofreram intervenção ou foram liquidados pelo Banco Central nos anos de 1995 e 1996 e por um grupo de bancos solventes como base de comparação. Foram selecionados para compor a amostra de bancos solventes 40 bancos

privados de grande porte. Cada um dos bancos detinha uma participação de seus ativos no total dos ativos do sistema bancário superior a 0,2% em 31.12.1994⁹.

Este critério de seleção da amostra do grupo de controle, formado pelas instituições solventes, foi escolhido por se tratar dos bancos de maior porte de nossa economia, sendo considerados os mais estáveis, tradicionais, de maior solidez, com maior diversificação de captação e melhor estrutura operacional no mercado. Este critério possibilita uma direta atualização do modelo ao longo do tempo.

3.2 Indicadores Econômico-Financeiros

O acompanhamento indireto (*off-site*) da situação financeira dos bancos geralmente é restrito a análise dos dados contábeis, o que pode ser alvo de críticas tanto pelo questionamento da veracidade dos valores informados como pelo fato de não refletir com exatidão os valores aceitos pelo mercado. No entanto, dados de mercado como preço de ações não estão disponíveis para a totalidade das instituições e, além disso, podem estar associados as companhias *holding* ao invés dos bancos. Portanto, para ser aplicável para a população de bancos e não somente a uma amostra, os sistemas de acompanhamento indireto (*early warning systems*), que tem por objetivo detectar com antecedência instituições com problemas, devem contar com os dados contábeis.

Os órgãos fiscalizadores dos Estados Unidos (*FDIC*, o *Federal Reserve Board* e o *OCC*) desenvolveram um sistema de classificação de bancos implantado em 1978 com a construção de indicadores econômico-financeiros que enfocavam cinco áreas genéricas: capital, qualidade dos ativos, gerenciamento, qualidade e montante de rentabilidade e liquidez. Estas áreas são identificadas pela palavra *CAMEL*¹⁰. Com o mesmo propósito do sistema *CAMEL*, de viabilizar avaliações da situação econômico-financeira das instituições e do seu conjunto, o Banco Central do Brasil também desenvolveu um sistema de acompanhamento das instituições financeiras sob seu controle, o sistema

⁹ A participação dos ativos das instituições financeiras no total dos ativos do sistema bancário foi extraído da revista Conjuntura Econômica da FGV.

¹⁰ CAMEL é uma sigla cujas letras significam **C**apital, **A**sset quality, **M**anagement, **E**arnings e **L**iquidity.

“INDCON - Indicadores Econômico-financeiros de Instituições sob Acompanhamento e Controle do Banco Central do Brasil”. Este sistema está apoiado numa estrutura de avaliação formada por sessenta e oito indicadores econômico-financeiros, classificados em cinco grupos gerais, estabelecidos em função da estrutura de capital, dos ativos, da rentabilidade, da eficiência gerencial e da liquidez dos bancos, o que deu origem à sigla “CAREL”.

Foram testados como possíveis previsores de insolvência bancária todos os indicadores econômico-financeiros que compõe a base de dados mantida pelo Banco Central do Brasil. Desta forma, ao selecionar indicadores utilizados pelo Departamento de Fiscalização do Banco Central, poderemos concluir se o modelo construído é uma ferramenta eficaz de previsão de insolvência de bancos no Brasil. Caso positivo, o modelo pode ser sugerido como um instrumento adicional à supervisão prudencial do sistema financeiro nacional.

3.3 Tratamento Estatístico

Como o tamanho da amostra de bancos é muito pequeno em comparação com o número de variáveis explicativas, torna-se necessário fazer uma seleção prévia dos indicadores econômico-financeiros, antes de incluí-los no modelo, com o intuito de aumentar os graus de liberdade, gerando assim um modelo mais parcimonioso.

O critério de exclusão de indicadores como possíveis previsores de insolvência bancária utilizado foi a aplicação do teste t sobre cada indicador. Primeiramente, calcula-se separadamente os valores médios dos indicadores econômico-financeiros das amostras de bancos solventes e insolventes. O teste t verifica se as médias dos indicadores das duas amostras são estatisticamente iguais, ao nível de significância de 5%, assumindo que ambas as amostras possuem a mesma variância. Ou seja, testamos $H_0 : \mu_1 = \mu_2$ versus $H_A : \mu_1 \neq \mu_2$, onde μ_1, μ_2 são as médias dos indicadores econômico-financeiros das amostras de bancos solventes e insolventes respectivamente. Se o nível de significância do teste é menor que 0.05, indica que as médias não são iguais

e devemos rejeitar H_0 . Este teste é feito para cada indicador. Quando o teste indica igualdade entre as médias é sinal de que o indicador não discrimina bem os dois grupos de bancos e, portanto, deve ser excluído da seleção.

Os indicadores selecionados no teste t entraram nos modelos como as variáveis explicativas. O procedimento “Stepwise” dentro da regressão logística e da regressão de Cox seleciona dentre os indicadores aqueles que mais contribuíram para a insolvência das instituições financeiras no período da análise e calcula a probabilidade de insolvência de cada banco. Caso esta probabilidade seja superior a 0.5, o modelo classifica o banco como insolvente, caso contrário como solvente. Comparamos esta classificação com a situação real do banco. Dois tipos de erros são possíveis. Como nos estudos anteriores sobre previsão de insolvência bancária, um **Erro Tipo I** é definido como a classificação errada pelo modelo de um banco que se tornou insolvente no período como sendo solvente. Um **Erro Tipo II** ocorre quando o banco é solvente e o modelo o classifica como insolvente. Quanto maior o percentual de acerto de classificação do modelo, maior a eficácia do modelo em prever o fenômeno de insolvência bancária.

3.4 Limitações dos Modelos

Os modelos de previsão de insolvência bancária desenvolvidos neste trabalho utilizaram-se de dados referentes à data-base de dezembro de 1994. A aplicabilidade destes modelos para outras datas-base precisará ser analisada em estudos futuros, já que existe a possibilidade de inclusão/exclusão de indicadores no modelo ou de mudança nos coeficientes dos indicadores estimados, devido à mudanças no ambiente econômico ao longo do tempo. Por se tratar de modelos *ex-post*, que necessitam da existência de um grupo de bancos solventes e outro de bancos insolventes, somente poderão ser recalculados para períodos mais recentes caso haja uma nova amostra de bancos insolventes.

4. Estimação do Modelo e Resultados

4.1 Teste t

O teste t para igualdade de médias entre os grupos de bancos solventes e insolventes foi aplicado aos 68 indicadores econômico-financeiros¹¹. A tabela 1 abaixo lista os 29 indicadores considerados significantes ao nível de 5%. Alguns destes indicadores apresentam alta correlação entre si e, portanto, excluiu-se um deles do modelo, já que a inclusão de ambos representaria “mais do mesmo”.

Tabela 1
Indicadores Econômico-financeiros Significantes ao nível de 5%

Capital

Participação dos recursos próprios nos recursos totais
Participação das captações no mercado interno nas exigibilidades
Participação das captações em moeda estrangeira nas exigibilidades
Participação de outros recursos nas exigibilidades
Comprometimento com operações ilíquidas

Ativos

Participação de ativos de realização imediata nas operações ativas usuais
Participação da carteira ativa no ativo total ajustado
Participação das operações ativas não usuais na carteira ativa
Participação das captações totais nas operações ativas usuais
Participação dos recursos sem remuneração nas operações ativas usuais
Participação dos recursos remunerados nas operações ativas usuais
Participação da carteira de câmbio
Taxa de inadimplência das operações de crédito
Nível de provisionamento dos ativos financeiros
Nível de provisionamento das operações de crédito

Rentabilidade

Rentabilidade mensal dos recursos próprios
Rentabilidade dos recursos próprios
Margem líquida
Taxa de retorno do ativo total ajustado
Margem operacional
Margem operacional média mensal no semestre
Remuneração média mensal do ativo operacional no semestre
Rentabilidade média mensal no semestre
Custo administrativo do ativo total ajustado

¹¹ O resultado do teste t para cada um dos 68 indicadores está descrito em Janot (1999).

Custo administrativo médio mensal do ativo total no semestre

Eficiência Gerencial

Nível de captação

Nível de aplicação

Nível de custo administrativo

Liquidez

Solvência geral

4.2 Modelo Logit

O modelo resultante da Análise de Regressão Logística, utilizando o método *Backward Stepwise* com a estatística *Wald*, selecionou apenas três entre os 26 indicadores utilizados na estimação como preditores de insolvência bancária, que são apresentados a seguir.

- Participação das captações em moeda estrangeira nas exigibilidades (CAPDOLAR)

Este indicador mensura a participação das captações em moedas estrangeiras em relação às exigibilidades. Uma tendência crescente indica um aumento do passivo em moeda estrangeira em relação às demais fontes de captação. A maior ou menor relevância deste indicador deve ser vista à luz de fatores conjunturais que recomendem ou não a concentração nas captações dessa modalidade. Como indica o resultado do modelo estimado, em dezembro de 1994, quanto maior a participação das captações em moeda estrangeira em relação as demais, menor a probabilidade de insolvência da instituição financeira. Este resultado é intuitivo, dado a conjuntura econômica vigente na época, onde vigorava um amplo diferencial entre as taxas de juros praticadas internamente e no exterior e uma desvalorização lenta e gradual da taxa de câmbio durante os dois anos subsequentes (prazo utilizado na classificação da amostra de bancos em solventes e insolventes).

- Participação das operações ativas não usuais na carteira ativa (PARTNUSU)

Este indicador apura a participação das operações ativas não usuais na carteira ativa. São consideradas não usuais aquelas operações não ligadas diretamente aos objetivos sociais da instituição, sendo que, uma tendência crescente revela maior direcionamento de recursos para essas aplicações. Podemos citar como operações ativas não usuais a aplicação de recursos em bens não destinados a uso, certificados de privatização, despesas antecipadas e outros créditos.

O sinal positivo do coeficiente estimado pelo modelo é intuitivo, demonstrando que quanto mais os bancos aplicarem em ativos não usuais, maior será a probabilidade de se tornarem insolventes. Em outras palavras, aplicações não previstas no objetivo social e nos regulamentos que regem as atividades do tipo da instituição representam riscos desaconselháveis.

- Participação de ativos de realização imediata nas operações ativas usuais (PARTATIM)

Este indicador mensura a participação dos ativos de realização imediata no total das operações ativas diretamente ligadas aos objetivos sociais da instituição. Uma tendência crescente indica uma concentração em ativos de liquidez imediata, o que diminui a probabilidade de insolvência. Este resultado foi encontrado pelo modelo, conforme esperado.

É interessante observar que nenhum indicador de rentabilidade, de eficiência gerencial e de liquidez foi considerado como bom preditor de insolvência pelo modelo. Dos três indicadores selecionados pelo modelo, dois estão relacionados a estrutura das aplicações (ativo) e um deles com a estrutura das origens de recursos (capital). Indicadores usualmente relevantes em estudos de insolvência bancária como custo administrativo e aqueles ligados ao comprometimento das aplicações, como taxa de inadimplência e nível de provisionamento das operações de crédito, não foram indicados pelo modelo

como significantes na previsão de insolvência bancária na conjuntura econômica brasileira do final de 1994.

O método *stepwise* adotado na estimação do modelo de regressão logística foi o de *backward selection*. Neste critério, todos os 26 indicadores selecionados como possíveis previsores de insolvência bancária entram no modelo na primeira etapa. Cada indicador é então analisado. Todos os que não atenderem ao critério de exclusão baseado em um *p-valor* (nível de significância) de 10% para a estatística selecionada são removidos. O próximo passo é reconsiderar as variáveis excluídas para inclusão. Quando nenhum indicador puder mais entrar ou ser removido do modelo, o algoritmo para.

A tabela 2 contém uma síntese dos resultados da estimação do modelo logit. Estão listados os três indicadores selecionados pelo modelo com seus respectivos coeficientes estimados, desvios padrões e níveis de significância.

TABELA 2
Resultados da Estimação do Modelo Logit

Indicador	Coefficiente	Desvio Padrão	Nível de Significância (<i>p-valor</i>)
CAPDOLAR	- 0,1897	0,0590	0,0013
PARTATIM	- 0,1071	0,0453	0,0181
PARTNUSU	0,4423	0,2058	0,0316
CONSTANTE	3,6200	1,6948	0,0327

O modelo de previsão de insolvência bancária construído classificou corretamente 91,8% dos bancos, correspondendo a 95% dos solventes e 85,71% dos insolventes, conforme mostra a tabela abaixo. A classificação correta ocorreu quando o banco era insolvente e tal fato foi previsto pelo modelo e quando o banco era solvente e o modelo o apontou como tal. O modelo classifica um banco como insolvente quando a probabilidade estimada de insolvência é maior do que 0.5.

TABELA 3
Classificação dos Bancos pelo Modelo Logit

Situação Real do Banco	Situação Estimada do Banco		Percentual de acerto de classificação
	SOLVENTE	INSOLVENTE	
SOLVENTE	38	2	95,00%
INSOLVENTE	3	18	85,71%

Percentual total de acerto de classificação pelo Modelo 91,80%

Dos três bancos insolventes classificados como solventes, conforme é apontado na tabela 3, dois deles se encontravam em uma faixa intermediária de insolvência (acima de 25%) e ficaram insolventes 15 e 16 meses após a data analisada, respectivamente. Geralmente, quando o banco se torna insolvente, a probabilidade de sobrevivência dele vai diminuindo a medida que a data de liquidação da instituição vai se aproximando¹². Para verificar esta afirmativa, apliquei o modelo estimado acima aos indicadores extraídos dos demonstrativos financeiros de dezembro de 1995 destes dois bancos. O modelo acertou em um deles, classificando-o como insolvente com probabilidade estimada de insolvência igual a 1.

4.2.1 Aplicação do modelo logit ao sistema bancário brasileiro

Como o modelo preditor de insolvência bancária apresentou um elevado percentual de acerto de classificação dentro da amostra, é necessário testar sua eficácia nos demais bancos do sistema financeiro. Efetuou-se sua aplicação aos dados de dezembro de 1994 dos 64 bancos solventes privados enquadrados como grande, médio e pequeno porte pelo sistema INDCON¹³. Não foram considerados os bancos de porte micro no teste, pois estes possuem certas especificidades que, dependendo do indicador selecionado como previsor de insolvência bancária, podem vir a deturpar a probabilidade estimada de insolvência. Excluíram-se também os bancos públicos, pois a decisão de liquidar ou intervir extrajudicialmente nestas instituições é muitas vezes uma decisão política. Isso

¹² Este fato pode ser observado melhor no modelo de risco proporcional de Cox, no qual a probabilidade de sobrevivência é uma função do tempo.

¹³ Vide Manual INDCON- Sistema de Indicadores Econômico-Financeiros de Instituições e Entidades sob Acompanhamento do Banco Central (1995).

leva o modelo a cometer o erro tipo II, ou seja, classificar como insolvente aqueles bancos que estão com a situação econômico-financeira debilitada, mas não foram liquidados no período analisado.

Um aceitável grau de acerto de classificação (75%) foi obtido. Este percentual subiria ainda mais se considerarmos que três dos 16 bancos classificados como insolventes vieram a ser liquidados posteriormente e outros 3 passaram por transferência de controle acionário. Esta é uma consideração importante, dado que existe uma defasagem temporal entre a insolvência econômica e a liquidação extrajudicial da instituição por parte do Banco Central. Portanto, a precisão do modelo é ainda melhor do que o indicado pelos resultados de classificação descritos na tabela 4 abaixo. Note que como estamos classificando somente bancos solventes, não existe a possibilidade de se cometer erro tipo I.

TABELA 4
Classificação dos bancos fora da amostra pelo modelo logit

	Frequência	Percentual	Percentual acumulado
Bem Clasificado	48	75,0	75,0
erro Tipo II	16	25,0	100,0
Total	64	100,0	

4.2.2 Validação do modelo ao longo do tempo

O grau de precisão resultante da aplicação do modelo estimado a todos os bancos, utilizando-se os dados contábeis de um e dois anos à frente, nos fornece uma importante medida do sucesso de um modelo de previsão de insolvência bancária. Como o modelo é construído com base em uma data passada e o seu objetivo é prever insolvências futuras, um elevado percentual de acerto na classificação dos bancos ao longo do tempo justificaria ainda mais sua adoção como um instrumento de supervisão do sistema financeiro nacional.

A tabela 5 mostra os resultados obtidos da aplicação do modelo logit para calcular a probabilidade de insolvência dos 231 bancos existentes no sistema financeiro em dezembro de 1995. Se a probabilidade estimada de insolvência de um determinado banco superar 0,5, o modelo está dizendo que o banco vai se tornar insolvente nos próximos dois anos a partir da data no qual os indicadores foram coletados.

TABELA 5
Classificação dos bancos em 1995

	Frequência	Percentual	Percentual Acumulado
Bem Classificado	<i>188</i>	<i>81,4</i>	<i>81,4</i>
Erro Tipo I	<i>8</i>	<i>3,5</i>	<i>84,8</i>
erro Tipo II	<i>35</i>	<i>15,2</i>	<i>100,0</i>
Total	<i>231</i>	<i>100,0</i>	

O alto percentual de acerto de classificação dos bancos em 1995 seria ainda maior se excluíssemos os bancos públicos e os privados de porte micro. Uma análise mais detalhada dos erros cometidos nos mostra que dentre os erros tipo II temos 17 bancos privados de porte micro, 7 bancos públicos, 2 bancos que faliram em 1998 e 2 bancos que tiveram o controle acionário transferido. Portanto, somente sete entre os 35 bancos listados como erro tipo II não se enquadravam nestes perfis. Com relação ao erro tipo I, dois dos 8 bancos classificados como solventes pelo modelo foram liquidados pelo Banco Central por prática de graves irregularidades na intermediação de compra e venda de títulos públicos de renda fixa e não por estarem com a situação financeira abalada, o que eleva ainda mais o percentual de acerto do modelo em 1995.

Os resultados obtidos da aplicação do modelo estimado, com base nos dados de 1994, para os indicadores econômico-financeiros de 223 bancos em dezembro de 1996 nos mostra que apesar de o percentual de acerto ter diminuído com relação ao ano anterior, ele ainda é expressivo, conforme mostra a tabela 6. Analogamente aos resultados referentes ao ano de 1995, o alto percentual de erro tipo II compõe-se de bancos públicos (14 erros), privados de porte micro (24 erros) e pelos bancos que transferiram o controle acionário (3 erros). Portanto, uma análise mais criteriosa dos erros nos mostra que o modelo pode ser aplicado até dois anos para frente de forma confiável. Entre os

bancos classificados como erros tipo I encontravam-se novamente os 2 bancos que foram liquidados por prática de irregularidades na intermediação de compra e venda de títulos públicos, o que não caracteriza erro de classificação do modelo.

TABELA 6
Classificação dos bancos em 1996

		Frequência	Percentual	Percentual Acumulado
	Bem Clasificado	157	70,4	70,4
	Erro Tipo I	11	4,9	75,3
	erro Tipo II	55	24,7	100,0
	Total	223	100,0	

O fato que o percentual de precisão não diminuiu muito ao longo do tempo mesmo sem a reestimação dos coeficientes do modelo é extremamente positivo. Isso indica que a relação entre as variáveis explicativas e as probabilidades de sobrevivência dos bancos, representados pelo modelo estimado, se mostrou relativamente estável no período. Esta é uma característica desejável de um modelo de previsão de insolvência bancária, já que isso previne a necessidade de atualizar os coeficientes do modelo ou mudar sua especificação freqüentemente. Cabe a ressalva de que mesmo com esta relativa estabilidade é importante ter sensibilidade econômica para saber quando houve uma mudança na conjuntura econômica que afete diretamente as variáveis preditoras de insolvência. Neste caso, é necessário a reestimação do modelo, para que este espelhe a nova realidade.

Os resultados da aplicação do modelo logit indicaram percentuais elevados de bancos bem classificados, sendo recomendável o sua utilização como uma ferramenta adicional na seleção das instituições financeiras que devem ser supervisionadas pelo Banco Central¹⁴.

¹⁴ Outras técnicas de avaliação da eficiência do ajuste deste modelo foram testadas, confirmando o bom ajustamento, e estão descritas em Janot (1999).

4.3 Modelo de risco proporcional de Cox

O modelo de risco proporcional de Cox, através do método *Forward Stepwise* com a estatística *Wald*, selecionou apenas três indicadores como os mais relevantes na previsão de insolvência bancária, que são apresentados a seguir:

- Custo administrativo médio mensal do ativo total no semestre (CUSADMED)

Este indicador mensura a participação do custo administrativo, inclusive encargos com pessoal, em relação ao ativo total médio da instituição, em termos de média mensal no semestre. O coeficiente estimado pelo modelo apresenta o sinal esperado, ou seja, um aumento dos encargos administrativos eleva a probabilidade de insolvência da instituição. Este indicador impacta diretamente na rentabilidade da instituição, pois representa um custo fixo que deverá ser acrescido à taxa de captação. Quanto maior for o indicador, maior será a taxa de aplicação viável para que o banco continue operando. No entanto, somente clientes de maior risco estarão dispostos a tomar emprestado a tais taxas.

- Participação de ativos de realização imediata nas operações ativas usuais (PARTATIM)
- Participação das captações em moeda estrangeira nas exigibilidades (CAPDOLAR)

Estes dois últimos indicadores também foram selecionados pelo modelo logit e, portanto, já foram analisados anteriormente.

O método *stepwise* adotado na estimação do modelo de regressão de Cox foi o de *forward selection*. Neste critério, o modelo começa como um modelo *baseline* sem qualquer variável nele. Os 26 indicadores são considerados um de cada vez e adicionados ao modelo se passarem no critério de seleção baseado em um *p-valor* de 5%. A medida que uma nova variável é adicionada ao modelo, as variáveis

anteriormente incluídas são avaliadas para exclusão. As que não forem significantes ao nível de significância de 10% são excluídas. Quando nenhum indicador puder mais ser adicionado ou removido, o algoritmo para.

A tabela 7 contém uma síntese dos resultados da estimação do modelo de Cox. Estão listados os três indicadores selecionados pelo modelo com seus respectivos coeficientes estimados, desvios padrões e níveis de significância.

TABELA 7
Resultados da Estimação do Modelo de Cox

Indicador	Coeficiente	Desvio Padrão	Nível de Significância (<i>p</i>-valor)
CAPDOLAR	- 0,0788	0,0306	0,0099
PARTATIM	- 0,0372	0,0156	0,0173
CUSADMED	0,6929	0,2125	0,0011

O modelo estimado acima pode ser usado para gerar a probabilidade que um banco sobreviva mais que t meses, aonde t pode variar de zero a 24 meses. Isso é feito pela substituição do valor do indicador econômico-financeiro relevante X do banco em análise, seu respectivo coeficiente estimado (B) e a probabilidade de sobrevivência *baseline* na equação $S(t / X, B) = S_0(t)^{\exp(X \cdot B)}$. Ao permitir t variar dentro de um intervalo, construímos o perfil de sobrevivência para o banco analisado.

A obtenção de um baixo percentual dos dois tipos de erros de classificação é o teste decisivo na validação de um potencial modelo de previsão de insolvência bancária. Os dois tipos de erros são importantes na avaliação da potencial utilidade de um modelo de *early warning*, no entanto, um bom modelo deveria exibir baixas taxas de erro tipo I. Não apontar uma insolvência com antecedência implica em altos custos, dado que inviabiliza qualquer medida corretiva por parte da instituição reguladora no sentido de evitar a liquidação e minimizar os custos. Entretanto, se o objetivo do modelo de previsão é ser útil na alocação de recursos escassos na fiscalização de bancos, as taxas

de erro tipo II também devem ser relativamente baixas. Existe uma exceção a esta regra geral. Como as previsões estão relacionadas ao horizonte de tempo especificado, alguns erros tipo II podem estar indicando bancos que vieram a quebrar em algum período de tempo futuro. Desta forma, um erro tipo II pode representar um sucesso do modelo, a medida que identifica um banco insolvente com uma antecedência ainda maior. Portanto, é de suma importância na avaliação da precisão do modelo que se identifiquem quais dentre os bancos classificados como insolventes pelo modelo foram liquidados pelo Banco Central posteriormente.

A probabilidade estimada de que um banco irá sobreviver mais do que t meses será comparada com o valor de corte 0.5. Se esta probabilidade for inferior a 0.5 o banco é classificado como insolvente pelo modelo. Caso contrário, o modelo o classifica como solvente. A análise apresentada aqui é centrada nas probabilidades de sobrevivência previstas para 6, 12, 18 e 24 meses. Por exemplo, um banco pode ter uma probabilidade estimada de sobreviver mais do que 6 meses igual a 0.7 e de sobreviver mais do que 12 meses igual a 0.4. Neste caso o modelo o classificará como solvente nos primeiros seis meses e como insolvente nos seis meses subsequentes.

Conforme definido anteriormente, um erro tipo I é a classificação errada de um banco que ficou insolvente dentro de algum intervalo específico de tempo durante o qual o modelo previu que não haveria a quebra e um erro tipo II é definido como a classificação errada de um banco que sobreviveu além do horizonte de tempo especificado e o modelo apontou a insolvência.

O modelo de Cox apresentou resultados satisfatórios com relação a classificação dos bancos dentro da amostra, conforme mostra a tabela 8. Na análise apresentada aqui, consideramos somente as probabilidades de sobrevivência previstas para 6, 12, 18 e 24 meses.

TABELA 8
Classificação dos bancos pelo modelo de Cox

Meses	Bem classificado	Erro tipo I	Erro tipo II
6	78,3% (47)	1,7% (1)	20% (12)
12	73,3% (44)	-	26,7% (16)
18	68,3% (41)	-	31,7% (19)
24	66,7% (40)	-	33,3% (20)

* A frequência dos bancos está computada entre parênteses.

No único caso de erro tipo I, o modelo o classificou como insolvente entre o sexto e o décimo segundo mês e sua liquidação ocorreu quatro meses após a data base de dezembro de 1994. Ou seja, o modelo errou o *timing* da insolvência, mas o classificou corretamente dentro do horizonte de tempo de 24 meses.

Com relação aos 12 bancos classificados como insolventes em 6 meses pelo modelo, 10 deles faliram entre o sexto e o décimo oitavo mês e os outros dois são os mesmos bancos classificados como erro tipo II pelo modelo logit dentro da amostra. Portanto, o resultado obtido para os primeiros 6 meses é um sucesso, tendo classificado corretamente 95% dos bancos, já que o modelo previu com uma antecedência ainda maior a insolvência de 10 instituições.

O resultado para os primeiros 12 meses também são bastante satisfatórios, dado que dos 16 bancos listados como erro tipo II, 6 deles faliram posteriormente, o que elevaria o percentual de acerto na classificação para 83,3%. No entanto, o grau de precisão do modelo decresce substancialmente quando o horizonte de tempo passa a ser 18 ou 24 meses.

4.3.1 Aplicação do modelo de risco proporcional de Cox ao sistema bancário brasileiro

O teste da eficácia do modelo aos demais bancos do sistema financeiro é necessário para validar o modelo. Efetuou-se sua aplicação aos dados de dezembro de 1994 dos mesmos 64 bancos solventes testados no modelo logit, obtendo-se o seguinte resultado:

TABELA 9
Classificação dos bancos fora da amostra pelo modelo de Cox

Meses	Bem classificado	Erro tipo I	Erro tipo II
6	98,4% (63)	-	1,6% (1)
12	92,2% (59)	-	7,8% (5)
18	81,3% (52)	-	18,8% (12)
24	81,3% (52)	-	18,8% (12)

* A frequência dos bancos está computada entre parênteses.

Este elevado percentual de acerto sobe para 100% nos primeiros 6 meses, considerando que o único erro tipo II passou por transferência de controle acionário em agosto de 1996 e sofreu intervenção em 1998. Com relação aos 12 erros tipo II obtidos para 24 meses, 5 deles ficaram insolventes nos dois anos subsequentes e 2 deles transferiram seu controle acionário, o que eleva o percentual de acerto para 92,2%.

4.3.2 Validação do modelo ao longo do tempo

A aplicação do modelo estimado para 231 bancos, utilizando-se os dados dos bancos um ano à frente (dezembro de 1995), produziu resultados extremamente positivos para os primeiros seis meses de previsão, onde dois dos 3 bancos classificados como solventes em 6 meses pelo modelo e sofreram intervenção pelo Banco Central, foram classificados como insolventes em 12 meses e, apenas dois dentre os 11 erros tipo II não eram bancos de micro porte nem públicos. Bancos com estas características representam a maior parte dos erros tipo II cometidos para os outros meses analisados. Por exemplo, para os 24 meses temos 31 bancos micro, 21 públicos e 7 que foram liquidados ou transferiram o controle acionário entre os erros tipo II. Além disso, os 2 bancos que foram liquidados por prática de irregularidades na intermediação de compra e venda de títulos públicos estão entre os erros tipo I como já era esperado.

A aplicação do modelo aos dados de dezembro de 1996 manteve os elevados percentuais de acerto, mostrando que o modelo é estável e válido ao longo do tempo¹⁵.

TABELA 10
Classificação dos bancos pelo modelo de Cox em 1995

Meses	Bem classificado	Erro tipo I	Erro tipo II
6	93,9% (217)	1,3% (3)	4,8% (11)
12	76,2% (176)	1,3% (3)	22,5% (52)
18	67,5% (156)	2,2% (5)	30,3% (70)
24	65,4% (151)	2,6% (6)	32% (74)

* A frequência dos bancos está computada entre parênteses.

TABELA 11
Classificação dos bancos pelo modelo de Cox em 1996

Meses	Bem classificado	Erro tipo I	Erro tipo II
6	95,1% (212)	3,1% (7)	1,8% (4)
12	75,3% (168)	3,1% (3)	21,5% (48)
18	66,4% (148)	2,2% (5)	31,4% (70)
24	65,5% (146)	3,1% (7)	31,4% (70)

* A frequência dos bancos está computada entre parênteses.

5. Conclusão

Este estudo mostrou que é possível identificar com antecedência as instituições financeiras mais propensas a se tornarem insolventes no Brasil, propiciando ao Departamento de Fiscalização do Banco Central a adoção de medidas corretivas em tempo hábil e uma alocação mais eficiente de seu quadro funcional para o acompanhamento direto das instituições.

Os dois modelos estatísticos analisados apresentaram percentuais elevados de acerto de classificação dos bancos, tanto dentro quanto fora da amostra, e mostraram-se válidos ao

¹⁵ Foram realizados diagnósticos dos modelos de Cox e da regressão logística não apontando observações problemáticas, o que indica que os modelos estão adequados. O procedimento adotado para fazer um diagnóstico é apresentado em Janot (1999).

longo do tempo, tendo previsto a insolvência da maior parte das instituições financeiras que foram liquidadas ou sofreram intervenção no período 1995/1998.

A insolvência bancária no Brasil ocorrida nos primeiros anos do Plano Real pode ser caracterizada por bancos com o seguinte perfil:

- baixa captação em moeda estrangeira em relação às demais fontes de captação;
- baixa participação de ativos de liquidez imediata no total das operações ativas usuais;

Estas características foram apontadas pelos dois modelos como fatores determinantes da insolvência bancária no período de 1995/1996. Além destes, o modelo logit apontou uma elevada aplicação de recursos em ativos que não estão diretamente ligados aos objetivos sociais da instituição como perfil de um banco insolvente. Já o modelo de risco proporcional de Cox indicou um custo administrativo elevado, inclusive encargos com pessoal, como fator relevante.

O modelo logit classificou corretamente 91,8% dos bancos inseridos na amostra e previu a insolvência de 18 entre as 21 instituições financeiras que compõe a amostra de bancos insolventes. Ao aplicar o modelo estimado aos bancos fora da amostra, o modelo teve um percentual de acerto de classificação igual a 75%. O modelo mostrou-se relativamente estável ao longo dos dois anos subsequentes a sua estimação, classificando corretamente mais de 70% da população de bancos.

O modelo de risco proporcional de Cox também apresentou elevados percentuais de acerto de classificação dos bancos, tendo ainda a vantagem, em relação ao modelo logit, de produzir não somente a probabilidade estimada de insolvência como também o tempo esperado para a quebra. A não observância de erro tipo I a partir do sexto mês valida o modelo de Cox como um bom modelo de previsão de insolvência bancária. No horizonte de seis meses, apenas 1 banco foi classificado como solvente e foi liquidado,

o que indica um percentual de erro tipo I de 1,7%. Ao apontar corretamente vinte das 21 instituições financeiras insolventes com antecedência, indicando ainda o período no qual a insolvência ocorreu, o modelo cumpre o papel de alertar o Banco Central sobre as futuras insolvências possibilitando ao mesmo a adoção de medidas saneadoras e corretivas que minimizem os custos ou até mesmo que evitem a liquidação das instituições debilitadas.

O modelo de Cox apresentou percentuais mais elevados de erro tipo II que o modelo logit na análise dos bancos dentro da amostra, resultando na classificação correta de 78,3% dos bancos em 6 meses; 73,3% em 12 meses; 68,3% em 18 meses; e, por último, 66,7% em 24 meses. Estes percentuais de acerto são inferiores ao obtido pelo modelo logit, mas devemos fazer uma ressalva. Como as previsões estão relacionadas ao horizonte de tempo especificado, alguns erros tipo II referem-se a bancos que vieram a ser liquidados em algum período de tempo futuro, o que resulta na previsão de insolvência destas instituições com uma antecedência superior a 6 meses, dando mais tempo para o Banco Central solucionar os problemas. Portanto, o erro tipo II representa um sucesso nesses casos. Considerando estes casos como acerto, o percentual de bancos bem classificados eleva-se substancialmente nos horizontes de tempo de seis meses (95%) e de doze meses (83,3%).

Os resultados obtidos da aplicação do modelo de risco proporcional de Cox aos demais bancos do sistema financeiro, mostrou que o modelo estimado é válido na classificação dos bancos fora da amostra, classificando corretamente 98,4% dos bancos em 6 meses; 92,2% em 12 meses e 81,3% em 18 e 24 meses. Estes percentuais são bem superiores ao obtido na aplicação do modelo logit.

Com relação a estabilidade do modelo de Cox estimado, observamos percentuais de acerto de classificação superiores a 90% para o horizonte de 6 meses, superiores a 75% para 12 meses e superiores a 65% para 18 e 24 meses utilizando-se os dados contábeis dos bancos de dezembro de 1995 e 1996. Assim como no modelo logit, estes percentuais se elevariam significativamente caso tivessem sido excluídos os bancos públicos e os bancos de porte micro.

Em lato senso, uma comparação dos resultados descritos acima indica que o modelo de risco proporcional de Cox é a técnica estatística mais adequada para prever o fenômeno de insolvência bancária no horizonte de um ano de antecedência, com níveis de acerto superiores a 90%. O modelo logit deve ser utilizado como um instrumento de supervisão no sentido de confirmar a classificação dos bancos pelo modelo de Cox, pois apesar de ter identificado uma proporção considerável das insolvências bancárias ocorridas no Brasil no período 1995/1998, não provê informação a respeito de quando dentro dos próximos dois anos a insolvência vai ocorrer.

Os resultados obtidos neste estudo demonstram que os demonstrativos financeiros dos bancos podem ser utilizados como fonte primária de dados na construção de modelos de previsão de insolvência bancária no Brasil.

Outros critérios de amostragem das instituições financeiras devem ser objeto de estudos futuros. A robustez do procedimento metodológico deve ser testada na construção de um modelo com um horizonte temporal de 12 meses. Nesse caso, teríamos uma amostra de bancos insolventes menor, mas, por outro lado, o modelo apresentaria uma defasagem temporal de apenas um ano, o que é um benefício na utilização prática dos modelos como instrumentos de supervisão prudencial do sistema financeiro. Também é possível estimar o modelo de risco proporcional de Cox desconsiderando a suposição de que os indicadores econômico-financeiros dos bancos permanecem constantes ao longo do período de análise¹⁶. No entanto, isto tornaria a análise bem mais complexa, devendo ser um tema de pesquisa adicional para a melhoria da precisão e confiabilidade dos modelos de previsão de insolvência bancária.

¹⁶ Na literatura internacional, este método é chamado de “Cox Regression with time-varying covariates”. Ver Norius (1993).

Referências

Amemiya, T. (1981), “Qualitative Response Models: A Survey”, *Journal of Economic Literature*, 19: 1483-1536.

Amemiya, T. (1994). *Introduction to statistics and econometrics*. Harvard University Press.

Andima (1997), “Reestruturação do SFN: uma análise do período 1994/1997” in *Retrospectiva 1997*.

Banco Central do Brasil (1995), “Manual INDCON - Sistema de Indicadores Econômico-Financeiros de Instituições e Entidades sob Acompanhamento do Banco Central”.

Barros, J. R.M. de e Mansueto F. de A. Jr. (1997), “Análise do Ajuste do Sistema Financeiro no Brasil”, Ministério da Fazenda, Secretaria de Política Econômica.

Cox, D.R. (1972), “Regression models and life-tables”, *Journal of the Royal Statistical Society, Series B*, 187-220.

Capelletto, L.(1995), “Contribuição à Avaliação do Desempenho Operacional de Bancos, Antes e Depois da Implantação do Plano Real”. Dissertação de Mestrado em Administração, Universidade de Brasília.

Cox, D.R. e D. Oakes (1984), *Analysis of survival data* (Chapman and Hall, London).

Demirguç-Kunt, A. (1989), “Deposit-Institution Failures: A Review of the Empirical Literature”, *Federal Reserve Bank of Cleveland Economic Review*, Fourth Quarter, 2-18.

Demirgüç-Kunt, A. e E. Detragiache (1997) “ The Determinants of Banking Crises: Evidence from Developing and Developed Countries” ”, IMF Working Paper, WP/97/106 (Washington: International Monetary Fund).

González-Hermosillo, B. (1996) “Banking Sector Fragility and Systemic Sources of Fragility”, IMF Working Paper, WP/96/12 (Washington: International Monetary Fund).

González-Hermosillo, B. (1996) “Banking System Fragility: Likelihood Versus Timing of Failure - An Application to the Mexican Financial Crisis”, IMF Working Paper, WP/96/142 (Washington: International Monetary Fund).

Greene, W.H. (1993) *Econometric Analysis*, Prentice Hall, Third Edition.

Janot, Marcio M. (1999), “Previsão de Insolvência Bancária no Brasil: aplicação de diferentes modelos entre 1995 e 1998”. Dissertação de Mestrado em Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Judge, G. G., Griffiths, W.E., Hill, C. Carter, Lutkepohl, Helmut, and Lee, Tsoung-Chao (1985), *The Theory and Practice of Econometrics*, New York: John Wiley & Sons, Inc.

Kiefer, N.M. (1988) “Economic Duration Data and Hazard Functions”, *Journal of Economic Literature*, 26: 646-79.

Lane, W.R., S.W. Looney, and J. W. Wansley (1986), “An Application of the Cox Proportional Hazards Model to Bank Failure”, *Journal of Banking and Finance*, 10: 511-31.

Lee, E. (1992) *Statistical Methods for Survival Data Analysis*, John Wiley & Sons, Inc..

Martin, D. (1977), “Early Warning of Banking Failure: A Logit Regression Approach”, *Journal of Banking and Finance*, 1: 249-76.

Matias, A.B. and J.O. Siqueira (1996), “Risco Bancário: Modelo de Previsão de Insolvência de Bancos no Brasil”, Revista de Administração, Abril/junho, 19-28.

Norius, M.J. (1993), “SPSS for Windows: advanced statistics”, release 6.0. Chicago. SPSS INC..

Rocha, Fabiana. (1999), “Previsão de Falência Bancária: um modelo de risco proporcional”, Pesquisa e Planejamento Econômico, pp.137-152

Rojas-Suárez, L. and Steven R. Weisbrod (1996) “Banking Crises in Latin America: Experiences and Issues.” In Hausmann, R. and Rojas-Suárez, L., eds., Banking Crises in Latin America, Inter-American Development Bank.

Rojas-Suárez, L. and Steven R. Weisbrod (1995) “Financial Market Fragilities in Latin America: the 1980s and 1990s, IMF Occasional Paper, No. 132 (Washington: International Monetary Fund).

Rojas-Suárez, L. (1998), “Early Warning Indicators of Banking Crises: What Works For Emerging Markets?”, Inter-American Development Bank.

Whalen, Gary (1991), “A Proportional Hazards Model of Bank Failure: An examination of Its Usefulness as na Early Warning Tool”, Federal Reserve Bank of Cleaveland Economic Review, First Quarter, 21-31.

Working Paper Series Banco Central do Brasil

1	Implementing Inflation Targeting in Brazil	Joel Bogdanski, Alexandre Antonio Tombini, and Sérgio Ribeiro da Costa Werlang	07/2000
2	Política Monetária e Supervisão do SFN no Banco Central	Eduardo Lundberg	07/2000
	Monetary Policy and Banking Supervision Functions on the Central Bank	Eduardo Lundberg	07/2000
3	Private Sector Participation: A Theoretical Justification of the Brazilian Position	Sérgio Ribeiro da Costa Werlang	07/2000
4	An Information Theory Approach to the Aggregation of Log-Linear Models	Pedro H. Albuquerque	07/2000
5	The Pass-through from Depreciation to Inflation: A Panel Study	Ilan Goldfajn and Sérgio Ribeiro da Costa Werlang	07/2000
6	Optimal Interest Rate Rules in Inflation Targeting Frameworks	José Alvaro Rodrigues Neto, Fabio Araújo, and Marta Baltar J. Moreira	09/2000
7	Leading Indicators of Inflation for Brazil	Marcelle Chauvet	09/2000
8	Standard Model for Interest Rate Market Risk	José Alvaro Rodrigues Neto	09/2000
9	Estimating Exchange Market Pressure and Intervention Activity	Emanuel-Werner Kohlscheen	11/2000
10	Análise do Financiamento Externo a Uma Pequena Economia	Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior	03/2001
11	A Note on the Efficient Estimation of Inflation in Brazil	Michael F. Bryan and Stephen G. Cecchetti	03/2001
12	A Test of Competition in Brazilian Banking	Márcio I. Nakane	03/2001
13	Modelos de Previsão de Insolvência Bancária no Brasil	Marcio Magalhães Janot	03/2001
14	Evaluating Core Inflation Measures for Brazil	Francisco Marcos Rodrigues Figueiredo	03/2001
15	Is It Worth Tracking Dollar/Real Implied Volatility ?	Sandro Canesso de Andrade and Benjamin Miranda Tabak	03/2001