

Emissão de cheques sem fundo: uma análise estatística*

Fernando Biagi da Silva**

Mestre em Economia pela Universidade
Católica de Brasília (UCB) e
Economista do Banco
do Brasil

Tito Belchior S. Moreira***

Doutor em Economia pela UnB e Professor
da Pós-Graduação de
Economia da UCB

Resumo

Este artigo utiliza um modelo classificatório — análise discriminante — para identificar um novo correntista como um emitente potencial, ou não, de cheques sem fundo. Os resultados mostram que é possível determinar o perfil daqueles que podem emitir cheques sem fundos com uma probabilidade de 86% de sucesso.

Palavras-chave

Cheques sem fundo; análise discriminante; classificação de crédito.

Abstract

This article uses a classificatory model—discriminant analysis—to identify a new client as a potential person who can, or not, write bounced checks. The results show that it is possible to determine the profile those that can wrote checks on non-existent balances with a probability of 86% of success.

* Artigo recebido em maio 2007 e aceito para publicação em maio 2009.

** E-mail: fernandobiagi@gmail.com

*** E-mail: tito@pos.ucb.br

Key words

Bounced checks; discriminant analysis; classification of credit.

Classificação JEL: C19, G21, G35.

1 Introdução

A decisão enfrentada pelos bancos comerciais de fornecer, ou não, um talão de cheques a um novo correntista é bastante similar à decisão que os mesmos devem tomar quando solicitados a conceder empréstimos. Em ambos os casos, o problema com que o banco se defronta é o de decidir, ao menos potencialmente, se está diante de um emissor de cheques sem fundos.

A literatura econômica tem avançado no caso de determinação de maus pagadores, também chamado de “problema de análise ou classificação de crédito”. Porém, quanto à emissão de cheques sem fundos, há pouca literatura acadêmica.

Para o processo de concessão de empréstimos, a literatura sugere o emprego do procedimento estatístico denominado “análise discriminante”. Como esse processo é similar ao problema de concessão de cheques, torna-se interessante investigar a contribuição que tal técnica pode trazer para a identificação de clientes que sejam emitentes de cheques sem fundos.

Ora, a emissão de cheques sem fundos traz custos para os bancos, pois seu produto cheque fica desacreditado. Aliás, esse é um produto bancário que necessita de confiança no mercado para que tenha boa aceitação. Por outro lado, pode-se intuir que a quantidade de cheques sem fundos passados nas lojas pode estar sendo alimentada por pessoas que têm acesso a um número muito grande de talões. Nesse contexto, tais pessoas podem tornar-se inadimplentes contumazes.

Vale ressaltar-se que, mesmo com a evolução dos meios de pagamentos eletrônicos (cartão de crédito, cartão de débito e transferências eletrônicas, dentre outros), o cheque continua tendo utilização. Esse fato decorre da transformação do cheque de ordem de pagamento à vista em um título de crédito, ou seja, o cheque pré-datado.

Sendo assim, torna-se interessante analisar a prática de emissão de cheques sem fundo no sentido de se identificarem situações que possam contribuir para valorizar esse meio de pagamento.

O objetivo deste estudo é empregar um modelo classificatório (utilizando análise discriminante) que possa identificar um novo correntista como um emissor potencial, ou não, de cheques sem fundo. Em outras palavras, pretende-se determinar o perfil daqueles clientes que podem emitir cheques sem fundos com base nas suas características cadastrais.

Este artigo baseia-se no trabalho desenvolvido por Barbosa, Faro e Araújo (1984), no qual foi aplicada a análise discriminante de dois grupos (aqueles que passaram cheques sem fundos e aqueles que nunca passaram), para determinar-se se as variáveis independentes idade, tempo de serviço e renda mensal poderiam contribuir para estabelecer uma função discriminante.

Além disso, não se utiliza a variável tempo de serviço, porém incluem-se as seguintes variáveis: estado onde se localiza a residência, estado civil, sexo, idade, grau de instrução, quantidade de dependentes, remuneração mensal, margem de contribuição do cliente, tempo que é cliente, residência própria e usuário de linha de crédito.

A amostra dessa pesquisa é fornecida por um banco comercial brasileiro com base em dados cadastrais de 7.033 clientes. Do total da amostra, 5.099 não emitiram cheques sem fundo, e 1.934 emitiram pelo menos um cheque sem fundo. Utiliza-se como variável dependente, ou variável de grupo, os dois tipos citados para o emprego da técnica estatística de análise multivariada — análise discriminante. Utiliza-se Tipo 0 para aqueles clientes que não emitiram cheques sem fundo, e Tipo 1 caso contrário. As variáveis explicativas que descrevem o perfil do cliente são extraídas dos dados cadastrais dos respectivos clientes (variáveis demográficas). Nesse contexto, a análise discriminante revela quais características dos clientes são estatisticamente significantes para distinguir os dois grupos: emissor e não emissor de cheques sem fundo.

O presente artigo está organizado da seguinte forma: após esta breve **Introdução**, faz-se a **Revisão da literatura**, apresentando-se os principais modelos sobre o tema. Na seção 3, desenvolve-se a metodologia empregada, e, por fim, nas seções 4 e 5, apresentam-se os **Resultados** e as **Considerações finais** respectivamente.

2 Revisão da literatura

Dentre os principais estudos elaborados sobre o tema “emissão de cheques sem fundos” conhecidos no Brasil, destacam-se os trabalhos de Barbosa, Faro e Araújo (1984), Souza (1985), Matos (1989) e Preece *et al.* (2004).

2.1 O estudo de Barbosa, Faro e Araújo

Um estudo importante que trata do problema de emissão de cheques sem fundos produzido por economistas brasileiros foi elaborado por Barbosa, Faro e Araújo (1984). Segundo esses pesquisadores, para efeito de análise econômica, podem-se considerar três tipos de indivíduos que emitem cheques sem fundos:

- a) o indivíduo que passa cheques sem fundos propositalmente;
- b) o indivíduo que, em meio a uma situação financeira difícil, envolvido por dívidas, utiliza-se desse meio para obter empréstimo. Assim, ele consegue melhorar a situação de liquidez de curto prazo. Esse cidadão pretende pagar a obrigação em um momento futuro;
- c) por fim, a falta de controle do saldo bancário, denominada descuido, leva o indivíduo a emitir cheques sem fundos.

O trabalho de Barbosa, Faro e Araújo (1984) leva em conta a situação do indivíduo que emite cheques sem fundos com o objetivo de obter crédito. Esse indivíduo pode adquirir bens e serviços, no primeiro momento, no valor total da sua renda, acrescido do montante de cheques sem fundos por ele emitidos, o que, de acordo com os autores, pode ser descrito do seguinte modo:

$$C_1 = r_1 + x \quad (1)$$

Onde: C_1 é igual ao total da compra de bens e serviços; r_1 é igual à renda no primeiro período; e x é igual ao volume de cheques sem fundos.

No segundo período, quando o indivíduo pretende saldar os cheques sem fundos, ele terá de deduzir da sua renda não somente o valor x dos cheques sem fundos, mas também terá de pagar uma penalidade p , imposta pela instituição financeira. Logo, a compra do segundo período será:

$$C_2 = r_2 - x - p \quad (2)$$

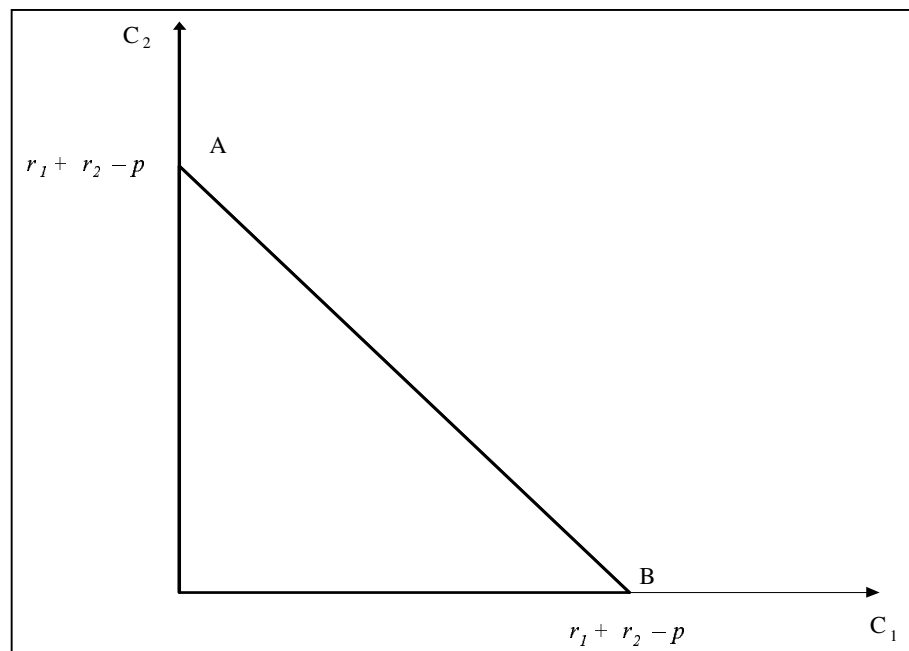
Onde r_2 é a renda que o indivíduo espera ter no segundo período. Substituindo a equação 1 na equação 2, obtém-se:

$$C_2 + C_1 = r_1 + r_2 - p \quad (3)$$

A Figura 1 representa graficamente a expressão 3: no eixo vertical, medem-se os gastos do segundo período, enquanto o eixo horizontal assinala os valores da compra de bens e serviços no primeiro período.

Figura 1

Restrição orçamentária do indivíduo



FONTE: Barbosa, Faro e Araújo (1984).

Analisando-se a Figura 1, fica fácil verificar que, ao aumentar a penalidade p , as opções de aquisição de bens e serviços do indivíduo diminuem, pois a rede AB desloca-se em direção à origem dos eixos, restringindo-se à área formada pela reta que toca os eixos.

Para os autores, a função utilidade pode ser descrita por:

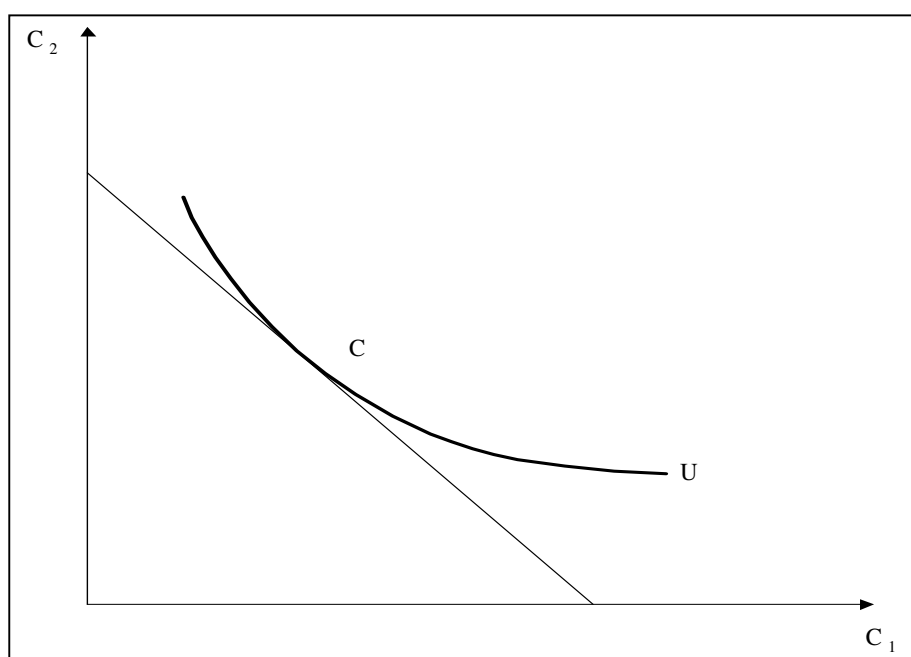
$$U(C_1; C_2) = U(C_1) + \delta U(C_2) \quad (4)$$

Onde δ é o fator de desconto, que depende das características do indivíduo.

Utilizando a hipótese de concavidade estrita e diferenciabilidade da função utilidade, ela pode ser representada graficamente através da Figura 2.

Figura 2

Função utilidade do indivíduo



FONTE: Barbosa, Faro e Araújo (1984).

O coeficiente angular da tangente a um ponto da curva de indiferença, como no ponto C, indica a razão pela qual o indivíduo está disposto a trocar o consumo nos dois períodos. Esse coeficiente, em valor absoluto, é igual a:

$$r = \frac{dC_1}{dC_2} = \frac{\partial U / \partial C_1}{\partial U / \partial C_2} \frac{1}{\delta} \quad (5)$$

Assim, o coeficiente angular é igual à razão entre as utilidades marginais dos gastos de consumo nos dois períodos, multiplicada pelo inverso do fator de desconto.

No estudo de Barbosa, Faro e Araújo (1984), a variável a ser determinada pelo agente econômico é o valor de x dos cheques sem fundos. Tal variável resulta do seguinte problema de maximização condicionada:

$$\text{Max } U(C_1; C_2) = U(C_1) + \delta U(C_2) \quad (6)$$

Sujeito à seguinte restrição:

$$C_2 + C_1 = r_1 + r_2 - p \quad (7)$$

A condição necessária e suficiente para um máximo interior é de que a razão entre as utilidades marginais seja igual ao fator de desconto:

$$\frac{\partial U / \partial C_1}{\partial U / \partial C_2} = \delta \quad (8)$$

Dessa forma, os autores afirmam que o equilíbrio será modificado, quando as condições objetivas com que se defronta o indivíduo, a renda do primeiro período, a renda esperada para o segundo período e a penalidade pelo uso da atividade ilegal mudarem.

Barbosa, Faro e Araújo (1984) propõem que a resolução conjunta das equações 1, 2 e 7 forneça uma equação para o valor dos cheques sem fundo, que, em geral, terá o seguinte formato:

$$x = x(r_1, r_2, p) \quad (9)$$

Onde x é uma função da renda r_1 do primeiro período, da renda r_2 do segundo período e da penalidade p . Para saber como x reage às mudanças de r_1 , r_2 e p , os autores defendem que é necessário fazer o seguinte exercício de estática comparativa:

Escreve-se a equação 10 como

$$\frac{\partial U}{\partial C_1} = \delta \frac{\partial U}{\partial C_2} \quad (10)$$

Diferenciando ambos os lados dessa expressão

$$dx = -\frac{U_{11}}{U_{11} + \delta U_{22}} dr_1 + \frac{\delta U_{22}}{U_{11} + \delta U_{22}} dr_2 - \frac{\delta U_{22}}{U_{11} + \delta U_{22}} dp \quad (11)$$

Onde o símbolo de $U_{ij}, i = 1, 2, \dots$ indica a derivada segunda, que é negativa, em virtude da suposição de que a função utilidade é côncava. Segue-se, então, dessa expressão que:

$$\begin{aligned} \frac{\partial x}{\partial r_1} &= -\frac{U_{11}}{U_{11} + \delta U_{22}} < 0 \\ \frac{\partial x}{\partial r_2} &= \frac{\delta U_{22}}{U_{11} + \delta U_{22}} > 0 \\ \frac{\partial x}{\partial p} &= \frac{\delta U_{22}}{U_{11} + \delta U_{22}} < 0 \end{aligned} \quad (12)$$

De acordo com os autores, as desigualdades equivalem às seguintes proposições:

- a) proposição 1 - o volume de cheques sem fundos diminui (aumenta), quando o nível de renda corrente aumenta (diminui);
- b) proposição 2 - o volume de cheques sem fundos aumenta (diminui), quando o nível de renda esperada aumenta (diminui);
- c) proposição 3 - o volume de cheques sem fundos diminui (aumenta), quando o nível de penalidade aumenta (diminui).

Também propõem, a título de ilustração, o seguinte caso:

$$U(C_1; C_2) = C_1^{1/2} + \delta C_2^{1/2} \quad (13)$$

A condição necessária e suficiente para a solução do problema de máximo condicionado é expressa por:

$$\frac{C_1^{1/2}}{C_2^{1/2}} = \delta \quad (14)$$

Utilizando as equações 1 e 2, resulta que o valor dos cheques sem fundos é dado por:

$$x = -\frac{\delta}{1+\delta^2}r_1 + \frac{1}{1+\delta^2}r_2 - \frac{1}{1+\delta^2}p \quad (15)$$

Assim, pode-se verificar que o coeficiente de r_1 é negativo, o de r_2 é positivo, e o coeficiente p é negativo, de acordo com as proposições.

Barbosa, Faro e Araújo (1984) afirmam que a equação 15 é importante, pois revela, de maneira bastante clara, uma faceta relevante do problema do cheque sem fundo. Desse modo, tais pesquisadores enfatizam que as características individuais desempenham papel fundamental nos coeficientes de reação das variáveis que afetam o volume de cheques sem fundos.

Declararam, ainda, que a penalidade imposta aos cheques sem fundos, variável controlada pelo Banco Central, influencia a quantidade de cheques sem fundos. Já a magnitude do impacto da penalidade sobre o valor dos cheques sem fundos depende das características individuais das pessoas que praticam essa ilegalidade.

Nesse sentido, há necessidade de se ter um conhecimento mais preciso das características individuais com relação ao uso do cheque.

Os referidos autores utilizaram a técnica estatística da análise discriminante para desenvolver um modelo econométrico com objetivo de identificar características individuais que determinem o praticante da atividade ilegal que é a emissão de cheque sem fundos.

Já os dados utilizados por eles foram fornecidos por dois bancos e continham as seguintes informações: cheques sem fundos (passou ou não passou), idade, tempo de serviço e remuneração mensal.

Barbosa, Faro e Araújo (1984) concluíram o estudo com as seguintes afirmações: (a) a análise estatística teve caráter exploratório face às dificuldades naturais que se encontram na obtenção de informações fidedignas sobre emissores de cheques sem fundos, mas a análise discriminante parece bastante promissora, pois revela um conjunto de informações úteis que contribuiria para aperfeiçoar a política de combate à emissão de cheques sem fundos; (b) deve ser estudada a possibilidade do estabelecimento de normas bancárias que requeiram um conjunto mínimo de informações qualitativas, bastante definidas, a respeito de cada correntista; (c) o poder de predição das funções discriminantes ficou ao redor de 75%; (d) as funções discriminantes apresentaram razoável poder classificatório, devido à baixa qualidade das informações levantadas, sugerindo desse modo, que essas não devem ser utilizadas para tomada de

decisão; (e) caso seja melhorada a qualidade da informação, a aplicação da técnica análise discriminante pode ser útil na prática.

2.2 O estudo de Souza

No estudo de Souza (1985), o problema de cheques sem fundos foi concebido sob o enfoque de oferta de delitos e, portanto, analisado pela teoria econômica tradicional, seguindo a formulação básica e pioneira proposta por Becker (1968).

Souza (1985) utilizou a abordagem da teoria da oferta de delitos e propôs um modelo explicativo do comportamento da quantidade de cheques sem fundos, no qual a variável é função inversa da probabilidade de punição, penalidade, retorno líquido em atividades legais ou legítimas, enquanto seria positivamente relacionada ao retorno líquido em atividades ilegais e à taxa de desemprego.

No estudo, foram incorporadas ao modelo variáveis ambientais ou específicas da realidade brasileira analisada (regulamentação e campanha de valorização do cheque), inclusive os custos e os ganhos de outras atividades ilegais que sejam substitutas ou complementares.

O autor construiu dois modelos econométricos que resultaram da aplicação da teoria da escolha ocupacional ao caso da prática de emissão de cheques sem fundos. O primeiro modelo considera dados mensais dos anos de 1981 a 1983, aplicando o modelo de equações simultâneas de dois estágios. Já o segundo modelo estima a função oferta de cheques sem fundos, considerando dados referentes ao período de janeiro de 1979 a dezembro de 1983.

As variáveis consideradas nesse estudo foram: relação entre o número mensal de cheques sem fundos e o número mensal de cheques compensados; probabilidade de punição, definida como a relação entre o número mensal de inclusões no CCF e o número mensal de cheques sem fundos; taxa de serviço paga na segunda apresentação do cheque; gastos do Banco Central, visando coibir a emissão de cheques sem fundos; valor real da arrecadação mensal de Impostos sob Circulação de Mercadoria (ICM), utilizada como *proxy* da renda real; taxas de juros mensais para aplicação de 180 dias; taxa efetiva mensal de *overnight* (atualmente taxa CDI); índice mensal das taxas de *overnight*, utilizado como *proxy* para índice mensal de rendimentos da atividade ilegal; índice mensal da taxa de arrecadação de ICM utilizado como *proxy* para índice mensal de rendimentos da atividade legal; relação entre o valor real médio estimado por cheque devolvido e o valor real da taxa de serviço; taxa de desemprego; *dummy* regulamentação; *dummy* campanha de valorização do cheque.

As conclusões a que o referido autor chegou foram as seguintes: (a) as variáveis utilizadas no modelo que são controladas pelo Banco Central se revelaram importantes na determinação da taxa observada de cheques sem fundos; (b) o diferencial entre os índices de rendimentos das atividades ilegais mostrou-se estatisticamente relevante; (c) as elasticidades estimadas das funções de oferta com respeito à penalidade foram superiores às elasticidades encontradas com respeito à probabilidade de punição, indicando aversão ao risco; (d) a probabilidade de punição mostrou ser fundamentalmente determinada pelo comportamento dos bancos, mais do que pelos gastos do Banco Central ou pela taxa de delitos.

2.3 O estudo de Matos

O estudo desenvolvido por Matos (1989) teve por objetivo identificar os fatores mais relevantes que determinam ou condicionam o comportamento dos índices de ocorrências de cheques sem fundos no Brasil como um todo e no Estado da Bahia em particular.

A equação 16 descreve o modelo formulado por Matos (1989) para a análise do comportamento do índice de ocorrências de cheques sem fundos, buscando determinar a repercussão de variáveis de natureza econômico-financeira e de medidas institucionais sobre o fenômeno.

$$I_t = b_0 + b_1 Y_t + b_2 H_t + b_3 J_t + b_4 M_t + b_5 C_t + b_6 L_t + u_t \quad (16)$$

- I_t é igual ao índice de cheque sem fundos, medido pela relação percentual entre a quantidade de cheques sem fundos e o total de cheques transitados pela câmara de compensação;
- Y_t é igual à taxa de variação do produto interno bruto (PIB) no trimestre t ;
- H_t é igual ao desemprego, aproximado pelo hiato do produto em $t-1$;
- J_t é igual à taxa de juros, aproximada pela taxa média real de financiamento do *overnight* em t ;
- M_t é igual ao valor real da multa aplicada ao emitente de cheques sem fundos em t ;
- C_t é igual à variável qualitativa que visa captar o efeito da campanha publicitária de valorização do cheque, veiculada na televisão nos períodos dez./80-fev./81, o qual assume valor 1; 4º trim./80-1º trim./81, 0; 1º trim./80-3º trim./80 e 1º trim./82-4º trim./88, 0,75; 0,50 e 0,25 em 2º trim./81-3º trim./81 e 4º trim./81, respectivamente;

- L_i é igual à variável *dummy* que visa captar o efeito do *law enforcement* por meio de medidas do Banco Central do Brasil, no sentido de inibir a prática de cheques sem fundos;
- u_i é igual ao termo aleatório.

O autor desenvolveu o estudo em dois planos. No primeiro, examinou a diferenciação dos índices de ocorrências de cheques sem fundos entre instituições financeiras. No segundo, analisou a evolução do fenômeno, procurando identificar os fatores conjunturais da economia ou as medidas governamentais que atuam sobre as ocorrências do delito.

Os resultados obtidos na primeira parte da análise foram: (a) as instituições financeiras estatais apresentaram índices de ocorrências de cheques sem fundos sensivelmente mais elevados do que suas congêneres controladas por capital nacional ou estrangeiro; (b) os bancos estrangeiros exibiram índices de cheques sem fundos muito mais baixos que os nacionais; e (c) não existiram evidências claras sobre a diferenciação de índices de cheques sem fundos por faixas de tamanho das instituições financeiras, mesmo quando os dados foram controlados pelo critério origem de capital. O autor observou que os pequenos bancos estatais apresentaram os índices mais elevados. Já no caso de bancos controlados pelo capital privado nacional ou estrangeiro, os níveis mais altos ficaram com os de maior porte.

Na segunda parte da análise, Matos (1989) chegou aos seguintes resultados: (a) os níveis da atividade econômica, medidos pela taxa de variação do PIB, afetam negativamente os índices de ocorrências de cheques sem fundos; (b) as taxas de desemprego num determinado trimestre afetam diretamente os índices de cheques sem fundos no período seguinte; (c) a implementação de novas regulamentações pelo Banco Central contribuiu significativamente para reduzir os índices de ocorrências de cheques sem fundos ao longo do tempo; (d) houve evidências que indicam que o valor da multa cobrada ao emitente de cheques sem fundos exerce efeito inibidor sobre a prática do delito; e (e) a campanha de valorização do cheque desenvolvida exerceu efeito significativo no sentido de reduzir a frequência relativa do fenômeno.

3 Metodologia

No presente estudo, é testada a aplicação da técnica estatística de análise discriminante com o objetivo de se verificar a aplicabilidade desse método, dado que, conforme concluíram Barbosa, Faro e Araújo (1984), havendo disponibilidade de dados com melhor qualidade, a técnica pode ser útil na prática.

Assim, neste estudo, é levantada uma base de dados sobre aproximadamente 7.000 indivíduos, onde são apuradas as seguintes informações: estado onde se localiza a residência; estado civil; sexo; idade; grau de instrução; quantidade de dependentes; remuneração mensal; margem de contribuição do cliente (o quanto de ganho a instituição financeira tem com o cliente); tempo que é cliente; residência própria; e se é usuário de linha de crédito.

A amostra foi fornecida por um banco comercial brasileiro, que não é revelado no presente estudo devido a possíveis efeitos de discriminações estatísticas. Na distribuição da base de clientes exposta na Tabela A.1, foram capturados dados cadastrais referentes à última semana do mês de novembro de 2004, num total de 7.033 novos clientes cadastrados em agências de todo o País. Do total da amostra, 5.099 não emitiram cheques sem fundo até o final do exercício de 2006, e 1.934 emitiram pelo menos um cheque sem fundo nesse mesmo período. Com base no pacote estatístico do SPSS, utilizam-se como variável dependente ou variável de grupo os dois tipos citados para o emprego da técnica estatística de análise multivariada — análise discriminante. Nesse contexto, utiliza-se Tipo 0 para aqueles clientes que não emitiram cheques sem fundo, e Tipo 1 caso contrário.

Conforme explicações de Hair *et al.* (2005), a regressão múltipla é, sem dúvida, a técnica de dependência multivariada mais utilizada, e isso se deve à habilidade que essa técnica tem em prever e explicar variáveis métricas. Porém, quando as variáveis não são métricas, a regressão múltipla não é adequada a essa questão. Os autores propõem o uso da análise discriminante, quando a variável dependente é não métrica.

Nessa situação, Hair *et al.* (2005) revelam que o pesquisador está interessado na previsão e na explicação das relações que provocam impacto na categoria em que um objeto está localizado.

Tabachnick e Fidell (1996) afirmam que o objetivo desse instrumental estatístico (análise discriminante) consiste em identificar as variáveis que melhor discriminam grupos previamente fixados. Em análise discriminante, os grupos são as variáveis dependentes (ou explicadas), enquanto as variáveis do problema representam as variáveis independentes (ou explicativas).

Malhotra (2004) destaca que os objetivos da análise discriminante são: estabelecer funções discriminantes, ou combinações lineares das variáveis independentes ou prognosticadoras, que melhor discriminem as categorias da variável dependente (grupos); verificar se existem diferenças significativas entre os grupos, em termos das variáveis prognosticadoras; determinar as variáveis preditoras que mais contribuem para as diferenças entre grupos; enquadrar ou classificar os casos em um dos grupos, com base nos valores das variáveis preditoras; avaliar a precisão da classificação.

As técnicas de análise discriminante, de acordo com Malhotra (2004), são definidas pelo número de categorias que a variável dependente possui. Neste estudo, a variável possui duas categorias (emissor ou não emissor de cheques sem fundos).

Hair *et al.* (2005) ressaltam que a análise discriminante é a técnica estatística apropriada para testar a hipótese de que as médias de grupo de um conjunto de variáveis independentes para dois ou mais grupos são iguais. Para tanto, a análise discriminante multiplica cada variável independente por seu peso correspondente e acrescenta esses produtos juntos. O resultado é um escore D discriminante composto para cada indivíduo da análise.

Ainda, conforme a explicação de Hair *et al.* (2005), calculando a média dos escores discriminantes para todos os indivíduos em um grupo, consegue-se a média do grupo, que é denominada centroide. Os centroides indicam o local mais típico de qualquer indivíduo de um grupo particular. Além disso, uma comparação dos centroides de grupos mostra quão afastados estão os grupos ao longo da dimensão testada.

Hair *et al.* (2005) salientam que o teste de significância estatística da função discriminante é uma medida generalizada da distância entre os centroides de grupos, ou seja, comparam-se as distribuições dos escores discriminantes para os grupos. Caso a sobreposição nas distribuições seja pequena, a função discriminante separa bem os grupos. No entanto, se a sobreposição é grande, a função é um discriminador pobre entre os grupos.

Alguns trabalhos com uso de análise discriminante podem ser destacados, como os de Dias, Moreira e Souza (2002), Moreira, Pinto e Souza (2004) e Bessa e Moreira (2007).

3.1 Análise discriminante

Essa técnica busca, basicamente, verificar se um elemento pertence a uma determinada categoria. Consiste num critério de alocação de novas observações a grupos previamente estabelecidos. O modelo tem sido amplamente utilizado em análises de risco de crédito e em análises de *performance* de instituições financeiras.

A variável dependente, ou variável estatística, é resultante da combinação linear de uma ou mais variáveis independentes. O valor previsto da função da discriminante é o escore discriminante, o qual é calculado para cada objeto (ou variável) sob análise.

A adequação da classificação das variáveis aos grupos predefinidos pode ser investigada adicionalmente por meio da função linear discriminante de Fischer, como citam Johnson e Wichern (1992).

Os principais objetivos da análise discriminante são: identificar as variáveis que melhor diferenciam ou separam grupos de indivíduos estruturalmente diferentes e mutuamente exclusivos; estimar os pesos ou os coeficientes da função discriminante; utilizar as variáveis e os coeficientes para estimar uma função discriminante que represente as formas entre os grupos ou as classes; utilizar a função estimada para classificar *a priori* novos indivíduos nos grupos.

A função discriminante tem a seguinte forma:

$$Z = \alpha + W_1 \cdot X_1 + W_2 \cdot X_2 + W_3 \cdot X_3 + \dots + W_n \cdot X_n \quad (18)$$

Onde Z é igual a escore discriminante ou categoria; α é igual a intercepto; W_i é igual a peso ou coeficiente discriminante para a variável i ; e X_i é igual à variável independente i .

As variáveis independentes relevantes, formadoras da função discriminante, são obtidas com base na estatística lambda (λ) de Wilks.

A estatística para cada uma das variáveis, segundo Maroco (2003), é dada por:

$$\lambda = \frac{SQE}{SQT} \quad \text{Sendo } SQE = \sum (y_i - \bar{y})^2 \text{ igual à soma dos quadrados dos}$$

erros dentro dos grupos; $SQT = \sum (y_i - \hat{y}_i)^2$ igual à soma dos quadrados totais; y igual à variável sob estudo; $\bar{y}$ igual a valor médio de y ; e $\hat{y}$ igual a valor estimado de y .

Os pesos ou coeficientes $w_{i1}, w_{i2}, \dots, w_{ip}$ são estimados de modo que a variabilidade dos escores (D_i) da função discriminante seja máxima entre os grupos e mínima dentro dos grupos. Eles representam uma medida relativa da importância das variáveis originais na função estimada. Quanto maior for o coeficiente das variáveis independentes, maior será a sua contribuição na discriminação entre os grupos, fator importante na interpretação e na análise do modelo.

Assim, dadas p variáveis e g grupos, é possível estabelecer $m = \min(g - 1; p)$ funções discriminantes que são combinações lineares das p variáveis, tal que:

$$D_i = w_{i1} \cdot X_1 + w_{i2} \cdot X_2 + \dots + w_{ip} \cdot X_p \quad (19)$$

com $i = 1, \dots, m$. O nível de significância da função é estimado com base em uma transformação qui-quadrado da estatística λ .

4 Resultados

Neste trabalho, foi aplicada a técnica estatística de análise classificatória para apurar uma função discriminante que caracterize os indivíduos que emitem cheques sem fundos e os que não emitem cheques sem fundos (para a análise dos dados, foi utilizado o *software* SPSS 11.5).

O primeiro passo da análise consiste em avaliar se as diferenças das médias entre os dois grupos de clientes são, ou não, significativas. Esses grupos são previamente determinados com base na amostra: um grupo corresponde aos clientes que nunca emitiram cheques sem fundos; o outro grupo é formado por aqueles que já emitiram cheques sem fundo. A ideia é saber quais as características dos clientes que distinguem os dois grupos. Por exemplo, pode-se verificar se as médias de idade dos clientes dos dois grupos são estatisticamente iguais. A hipótese nula é que as médias de idades são estatisticamente iguais. Com base no teste F, na Tabela 1, pode-se ver que não se aceita H_0 . Nesse caso, a idade é uma das variáveis que diferencia os clientes dos dois grupos.

A Tabela 1 apresenta os resultados dos testes de média dos dois grupos. Verifica-se que, em um nível de significância de 5%, existem diferenças significativas em nove variáveis, tendo sido excluídas da análise subjacente as variáveis estado civil e sexo. Nesse contexto, o estado civil e o sexo dos clientes não são características que diferenciam os emissores dos não emissores de cheques sem fundos.

Como há dois grupos, estima-se apenas uma função discriminante. O autovalor associado a essa função é 0,934 e responde por 100% da variância explicada. A correlação canônica associada a essa função é 0,695 (Tabela 2). O quadrado dessa correlação, $(0,695)^2 = 0,48$, indica que 48% da variância na variável dependente são explicados por esse modelo.

Tabela 1

Testes de igualdade das médias dos grupos

VARIÁVEIS	WILKS' LAMBDA	F	Df1	df2	Sig.
Zscore: (UF)	0,998	11,939	1	7 031	0,001
Zscore: estado civil	1,000	0,587	1	7 031	(1)0,444
Zscore: (sexo)	1,000	3,219	1	7 031	(1)0,073
Zscore: (idade)	0,989	78,958	1	7 031	0,000
Zscore: dependentes	0,986	96,479	1	7 031	0,000
Zscore: renda mensal	0,981	134,504	1	7 031	0,000
Zscore: escolaridade.....	0,995	36,896	1	7 031	0,000
Zscore: margem de contri- buição total	0,984	113,119	1	7 031	0,000
Zscore: tempo de cliente ...	0,999	6,522	1	7 031	0,011
Zscore: linha de crédito	0,993	47,799	1	7 031	0,000
Zscore: residência própria	0,553	5 689,893	1	7 031	0,000

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

(1) Não significativa.

Tabela 2

Autovalor

FUNÇÃO AUTOVALOR	PERCENTUAL DA VARIÂNCIA	PERCENTUAL ACUMULADO	CORRELAÇÃO CANÔNICA
1	0,934	100,0	0,695

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

Para comprovar que os dois grupos são estatisticamente diferentes, utiliza-se o valor do teste do Wilks' Lambda.

Segundo Hair *et al.* (2005), o Wilks' Lambda pode ser transformado numa variável com uma distribuição aproximada de χ^2 (qui-quadrado), podendo, então, ser utilizado para determinar se os dois grupos são estatisticamente diferentes.

O valor do Lambda representa a proporção da variância total dos coeficientes discriminantes não explicada pela diferença entre os grupos.

Como se pode observar na Tabela 3, o valor do Lambda, calculado em 0,517, foi transformado num valor do qui-quadrado de 4.634,825, com 11 graus de liberdade e estatisticamente significativo em um nível de 1%. Portanto, parece pouco provável que os clientes emitentes e os não emitentes de cheques sem fundos tenham as mesmas médias na função discriminante.

Tabela 3

Coeficientes da função discriminante padronizada

TESTE DE FUNÇÕES	WILKS' LAMBDA	QUI-QUADRADO	g.l.	Sig.
1	0,517	4 634,825	11	0,000

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

Dada a significância estatística entre os dois grupos, é conveniente examinar-se a contribuição individual das variáveis na função discriminante. Os coeficientes das variáveis representam em valor absoluto a importância relativa de cada uma, para distinguir entre clientes emitentes e não emitentes de cheques sem fundos.

Como ressaltado, o objetivo da análise é determinar quais variáveis são mais eficientes para a discriminação de clientes que podem emitir cheques sem fundo e, posteriormente, elaborar uma função discriminante. Destarte, o propósito da análise discriminante é definir o conjunto de variáveis que discriminará melhor os grupos que apresentam grande valia.

Percebe-se que a variável residência própria é o prognosticador mais relevante na discriminação entre os grupos, seguida por renda mensal, margem de contribuição, dependentes e idade (Tabela 4).

Pode-se ter ideia da importância relativa dos prognosticadores examinando as correlações estruturais, também denominadas cargas canônicas ou cargas discriminativas. Essas correlações simples entre cada prognosticador e a função discriminante representam a variância que o preditor compartilha com a função.

Malhotra (2004) destaca que essas correlações devem ser interpretadas com cautela. Eis que, ao se examinar com cuidado a Tabela 5, percebe-se que a mesma sugere algumas características ao emissor de cheques sem fundo: os emissores desses cheques são pessoas mais jovens, que possuem dependentes, fornecem maiores margens de contribuição, têm residência própria, utilizam linhas de crédito e possuem maior grau de instrução e remuneração mensal.

Tabela 4

Função discriminante estimada na média dos grupos

VARIÁVEIS	FUNÇÃO 1
Zscore: residência Própria	0,931
Zscore: renda mensal	0,143
Zscore: margem de contribuição total	0,131
Zscore: dependentes	0,121
Zscore: (idade)	0,110
Zscore: linha de crédito	0,085
Zscore: escolaridade	0,075
Zscore: (UF)	-0,043
Zscore: tempo de cliente	0,032
Zscore: (sexo)	0,022
Zscore: estado civil	0,009

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

Tabela 5

Coeficientes da função classificação

VARIÁVEIS	EMISSOR	
	Tipo 0	Tipo1
Zscore: (UF)	0,030	-0,080
Zscore: estado civil	-0,025	0,066
Zscore: (sexo)	0,017	-0,045
Zscore: (idade)	0,117	-0,310
Zscore: dependentes	-0,064	0,169
Zscore: renda mensal	-0,062	0,164
Zscore: escolaridade	-0,028	0,074
Zscore: margem de contribuição total	-0,095	0,250
Zscore: tempo de cliente	-0,091	0,240
Zscore: linha de crédito	-0,051	0,135
Zscore: residência própria	-0,785	2,069
Constante	-0,870	-1,924

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

NOTA: Função Discriminante Linear de Fisher.

Com base na equação 17 e na Tabela 5, pode-se explicitar a função discriminante:

$$D = -1,924 + 2,069 * (res.própria) - 0,310 * (idade) + 0,250 * (MC) \\ + 0,240 * (tempode cliente) + 0,169 * (dependentes) + 0,164 * (renda) \\ + 0,135 * (crédito) - 0,080(UF) + 0,074 * (nível de escolaridade) \\ + 0,066 * (estadocivil) - 0,045 * (sexo)$$

A função discriminante, explicitada acima, mostra as variáveis independentes ponderadas pelos seus respectivos pesos em ordem decrescente. Observa-se, inicialmente, que as variáveis ponderadas com os menores pesos são estado civil e sexo, o que confirma o resultado apresentado na Tabela 1. Outros resultados interessantes são apresentados a seguir:

- a) os clientes mais jovens são mais propensos a emitirem cheques sem fundos, o que é um resultado esperado, se for considerado o argumento de Barbosa, Faro e Araújo (1984)¹. Assim, considera-se que indivíduos mais jovens têm menor controle do saldo bancário e são mais descuidados, o que os levaria a emitirem cheques sem fundos;
- b) os clientes que têm acesso a linhas de crédito são mais propensos a emitirem cheques sem fundos. Pois, tomando-se o acesso a linhas de crédito como uma *proxy* para endividamento, pode-se utilizar o argumento de Barbosa, Faro e Araújo (1984)², no qual indivíduos endividados, com situação financeira difícil, são mais propensos a emitirem cheques sem fundos. Nesse contexto, trata-se de um resultado esperado. Numa linha de argumento similar, pode-se pensar que indivíduos com dependentes provavelmente possuem uma situação financeira mais difícil, com menor liquidez, o que poderia explicar uma maior propensão à emissão de cheques sem fundos;
- c) segundo Matos (1989), os níveis de atividade econômica (níveis de renda) têm uma relação inversa com os índices de ocorrências de cheques sem fundos, com base em dados agregados. Os presentes resultados, com base em microdados, mostram justamente o contrário. Tais resultados se confirmam através da análise de várias *proxies* para o nível de renda, como, por exemplo, a própria variável renda mensal e

¹ Ver item c do primeiro parágrafo da seção 2.1.

² Ver item b do primeiro parágrafo da seção 2.1.

outras que são indicadoras de clientes com maior poder aquisitivo, tais como: ter residência própria, ter uma maior margem de contribuição, maior nível educacional (*proxy* para maior capital humano e, portanto, para maior rendimento). Os resultados mostram que, quanto maior for o padrão de vida dos clientes, maior será a propensão à emissão de cheques sem fundos.

Uma possível explicação para os resultados apresentados no item c, decorre do fato de que indivíduos de baixa renda podem ser mais avessos ao risco, dada a pouca capacidade de acesso ao crédito, seja via empréstimos, seja via utilização de cheques pré-datados. O custo de ter o “nome sujo na praça” é muito maior do que para aqueles indivíduos que possuem maiores níveis de renda. Nessa mesma linha de raciocínio, aqueles clientes que possuem um menor tempo como correntistas podem ser mais avessos ao risco, pois ou ainda não possuem cheque especial, ou, se possuem, têm menores limites para esses cheques. O mesmo também se aplica ao acesso ao crédito. Portanto, o custo para o novo correntista de emitir cheques sem fundos, em princípio, deve ser maior do que para os clientes antigos.

Por fim, para testar a validade da função discriminante obtida, é necessário determinar-se se os indivíduos previamente classificados nos respectivos grupos diferem significativamente da classificação desses mesmos indivíduos com base na função discriminante obtida. Os resultados da validação obtidos do cruzamento da função discriminante da amostra utilizada são apresentados na Tabela 6. Dessa validação, verifica-se que 86% dos indivíduos foram corretamente classificados nos grupos por meio da função discriminante.

A partir dos resultados, é possível concluir-se que as variáveis significantes aqui utilizadas são úteis para explicar o comportamento dos clientes face à possibilidade de emitir, ou não, cheques sem fundos, com uma percentagem de acerto, em média, de 86%.

Tabela 6

Resultado da classificação

DISCRIMINAÇÃO	GRUPO PREDITO		
	Tipo 0 (1)	Tipo 1 (2)	Total
Original			
Grupo 0	4 451	648	5 099
Grupo 1	340	1 594	1 934
Percentual			
Grupo 0	87,3	12,7	100,0
Grupo 1	17,6	82,4	100,0
Validação cruzada			
Grupo 0	4 450	649	5 099
Grupo 1	340	1 594	1 934
Percentual			
Grupo 0	87,3	12,7	100,0
Grupo 1	17,6	82,4	100,0

FONTE DOS DADOS BRUTOS: Pesquisa de campo.

NOTA: 1. Estão corretamente classificados 86,0% dos indivíduos.

2. Estão corretamente classificados pelo procedimento de validação cruzada 85,9% dos indivíduos.

(1) Não emitiu cheques sem fundo. (2) Emitiu cheques sem fundo.

5 Considerações finais

A decisão enfrentada pelos bancos comerciais de fornecer, ou não, um talão de cheques a um novo correntista é bastante similar à decisão que os mesmos devem tomar quando solicitados a conceder empréstimos. Em ambos os casos, o problema com que o banco se defronta é o de decidir se, ao menos potencialmente, está diante de um emitente, ou não, de cheques sem fundos.

O objetivo deste estudo é apurar um modelo classificatório (utilizando análise discriminante) que possa identificar um novo correntista como um emitente potencial, ou não, de cheques sem fundo.

A pesquisa possui por base o trabalho desenvolvido por Barbosa, Faro e Araújo (1984), no qual foi aplicada análise discriminante de dois grupos (aqueles que passaram cheques sem fundos e aqueles que nunca passaram), para deter-

minar se as variáveis independentes idade, tempo de serviço e renda mensal poderiam contribuir para estabelecer uma função discriminante.

Além das variáveis independentes idade e renda mensal, também foram incluídas as seguintes informações: estado onde se localiza a residência, estado civil, sexo, grau de instrução, dependentes, margem de contribuição do cliente, tempo que é cliente, se possui residência própria e se o mesmo é usuário de linha de crédito.

De acordo com os resultados, conclui-se que as variáveis significantes são úteis para determinar o comportamento dos clientes face à possibilidade de emitirem, ou não, cheques sem fundos. Os resultados sugerem que o emissor de cheques sem fundos possui o seguinte perfil: pessoa mais jovem, possui dependentes, fornece maior margem de contribuição, tem residência própria, utiliza linhas de crédito e possui maior grau de instrução e remuneração mensal.

Um resultado que merece destaque é a relação direta entre níveis de renda e ocorrências de cheques sem fundos. Os resultados mostram que, quanto maior for o padrão de vida dos clientes, maior será a propensão à emissão de cheques sem fundos. Uma possível explicação para esse resultado decorre do fato de que indivíduos de baixa renda podem ser mais avessos ao risco, dada a pouca capacidade de acesso ao crédito, seja via empréstimos, seja via utilização de cheques pré-datados. O custo de ter o “nome sujo na praça” é muito maior comparativamente aos indivíduos que possuem maiores níveis de renda.

O argumento acima apresentado sugere pesquisas com base nas seguintes perguntas: (a) quem são os maiores devedores ou “inadimplentes” da União ou das instituições financeiras: os grandes municípios ou os pequenos? Os grandes estados da Federação ou os pequenos? Os grandes produtores rurais ou os pequenos? As grandes empresas ou as pequenas? Pode-se sugerir também, para futuras pesquisas, um método de validação dos resultados da análise discriminante; dado que já se conhecem as variáveis de grupo previamente, com base apenas nas variáveis independentes ou explicativas, poder-se-ia sugerir separar os indivíduos em grupos, utilizando a análise de *cluster*, também conhecida como análise de agrupamento. Isto posto, poderia verificar-se se realmente os grupos de indivíduos determinados pela análise de *cluster* correspondem aos grupos previamente definidos.

Destaque-se que, no estudo desenvolvido por Barbosa, Faro e Araújo (1984), o poder de predição das funções discriminantes ficou ao redor de 75%, enquanto, no presente trabalho, a percentagem de indivíduos corretamente classificados corresponde a 86%.

Anexo

Tabela A.1

Distribuição dos clientes, por unidades da Federação, no Brasil — nov./04

CÓDIGOS	ESTADOS	QUANTIDADE DE INDIVÍDUOS
1	AC	19
2	AL	110
3	AM	63
4	AP	18
5	BA	515
6	CE	192
7	DF	263
8	ES	104
9	GO	230
10	MA	163
11	MG	818
12	MS	121
13	MT	221
14	PA	165
15	PB	176
16	PE	270
17	PI	106
18	PR	497
19	RJ	662
20	RN	160
21	RO	28
22	RR	37
23	RS	483
24	SC	346
25	SE	67
26	SP	1 149
27	TO	50
TOTAL		7 033

FONTE: Pesquisa de campo.

Referências

- BARBOSA, Fernando de H.; FARO, Clovis; ARAUJO, Aloísio P. **Uma análise estatística das causas da emissão do cheque sem fundos**: formulação de um projeto piloto. Rio de Janeiro: FGV, 1984. (Ensaio Econômico EPGE, n. 54).
- BECKER, Gary. Crime and punishment: an economic approach. **Journal of Political Economy**, n. 76, 1968.
- BESSA, Cristina F. M.; MOREIRA, Tito B. S.; PINTO, Maurício Barata de P. Fatores comuns determinantes de crises cambiais e financeiras. (Working papers, n. 07-006). Disponível em: <www.ucb.br>. Acesso em: set. 2007.
- DIAS, José.; MOREIRA, Tito B. S.; SOUZA, Geraldo da S. Fatores determinantes de crises cambiais e bancárias. **Revista Econômica do Nordeste**, v. 33, n. 2, abr.-jun., 2002.
- HAIR, J. et al. **Análise multivariada de dados**. Porto Alegre: Bookman, 2005.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 3. ed. New Jersey: Prentice Hall, 1992.
- MALHOTRA, Naresh K. **Pesquisa de marketing**: uma orientação aplicada. Porto Alegre: Bookman, 2004.
- MAROCO, J. **Análise estatística com utilização do SPSS**. 2. ed. Lisboa: Silabo, 2003. p. 56-59, p. 112-114 e p. 332-360.
- MATOS, Orlando Carneiro de. **Fatores determinantes da prática de cheques sem fundos, 1980/88**. Salvador: Banco Central do Brasil, 1989.
- MOREIRA, T. B. S.; PINTO, M. B. de P.; SOUZA, G. S. Uma metodologia alternativa para mensuração de pressão sobre o mercado de câmbio. **Estudos Econômicos**, v. 34, n. 1, p. 73-100, jan./mar., 2004.
- PREECE, D. et al. Agency theory and the house bank affair. **Review of Financial Economics**, v. 13, n. 3, 2004.
- SOUZA, Ubiratan Jorge Iório de. **Uma análise econômica do problema cheque sem fundos no Brasil**. Rio de Janeiro: IBMEC, 1985.
- TABACHNICK, B.; FIDELL, L. **Using multivariate statistics**. 3. ed. New York: Harper Collins, 1996.

