

DOI: 10.7819/rbgn.v16i52.1555

ÁREA TEMÁTICA: CONTABILIDADE E CONTROLADORIA

Fraudes Contábeis: uma estimativa da probabilidade de detecção

Accounting Fraud: an estimation of detection probability

Fraudes contables: una estimación de la probabilidad de detección

Artur Filipe Ewald Wuerges¹
José Alonso Borba²

Recebido em 20 de fevereiro de 2013 / Aprovado em 01 de outubro de 2014

Editor responsável: Ivam Ricardo Peleias, Dr.

Processo de avaliação: *Double Blind Review*

RESUMO

Fraudes nas demonstrações financeiras (FDF) custam caro para os investidores e podem prejudicar a credibilidade dos auditores. Para prevenir e detectar fraudes, é útil conhecer suas causas. Os modelos de escolha binária (por exemplo, logit e probit), frequentemente utilizados na literatura, porém, não levam em consideração os casos de fraudes não detectados e, portanto, apresentam testes de hipóteses pouco confiáveis. Usando uma amostra de 118 empresas acusadas de fraude pela Comissão de Valores Mobiliários dos Estados Unidos (Securities and Exchange Commission, SEC), estimamos um modelo logit que corrige os problemas oriundos de fraudes não detectadas em empresas dos Estados Unidos. Para evitar problemas de multicolinearidade, extraímos sete fatores a partir de 28 variáveis, usando o método dos componentes principais. Nossos resultados indicam que apenas 1,43% dos casos de FDF foram divulgados pela SEC. Das sete variáveis significativas incluídas em um modelo logit tradicional e não corrigido, três na realidade não

foram consideradas significativas em um modelo corrigido. A probabilidade de FDF é 5,12 vezes maior quando o auditor da empresa emite um parecer adverso ou com ressalvas.

Palavras-chave: Fraude contábil. AAER. Erros de classificação. Logit. Análise fatorial.

ABSTRACT

Financial statement fraud (FSF) is costly for investors and can damage the credibility of the audit profession. To prevent and detect fraud, it is helpful to know its causes. The binary choice models (e.g. logit and probit) commonly used in the extant literature, however, fail to account for undetected cases of fraud and thus present unreliable hypotheses tests. Using a sample of 118 companies accused of fraud by the Securities and Exchange Commission (SEC), we estimated a logit model that corrects the problems arising from undetected frauds in U.S. companies. To avoid multicollinearity problems, we extracted seven factors from 28 variables using the principal

1. Mestre em Administração pela Universidade Federal de Santa Catarina (UFSC). [artur.wuerges@uffs.edu.br]

2. Doutor em Contabilidade pela Universidade de São Paulo (USP). [jalonso@cse.ufsc.br]

Endereço dos autores: Rua Major Antônio Cardoso, 590, CEP: 97900-000 – Cerro Largo – RS – Brasil

factors method. Our results indicate that only 1.43 percent of the instances of FSF were publicized by the SEC. Of the six significant variables included in the traditional, uncorrected logit model, three were found to be actually non-significant in the corrected model. The likelihood of FSF is 5.12 times higher when the firm's auditor issues an adverse or qualified report.

Keywords: Accounting fraud. AAER. Misclassification. Logit. Factor analysis.

RESUMEN

El fraude en los estados financieros (FEF) es costoso para los inversionistas y pueden minar la credibilidad de los auditores. A fin de prevenir y detectar el fraude, es útil conocer sus causas. Sin embargo, los modelos de elección binaria (logit y probit, por ejemplo) a menudo utilizados en la literatura, no tienen en cuenta los casos de fraudes detectados y consecuentemente presentan pruebas de hipótesis poco fiables. Utilizando una muestra de 118 compañías acusadas de fraude por la Comisión de Bolsa y Valores de EE.UU. (Securities and Exchange Commission, SEC), hemos estimado un modelo logit que corrige los problemas derivados de los fraudes no detectados en las compañías estadounidenses. Para evitar problemas de multicolinealidad, hemos extraído siete factores de 28 variables, utilizando el método de componentes principales. Nuestros resultados indican que sólo el 1,43% de los casos de FEF se han dado a conocer por la SEC. De las siete variables significativas incluidas en un modelo logit tradicional y no corregido, tres en efecto no fueron consideradas significativas en un modelo corregido. La probabilidad de FEF es 5,12 veces mayor cuando el auditor de la compañía emite una opinión adversa o con reservas.

Palabras clave: Fraude contable. AAER. Errores de clasificación. Logit. Análisis factorial.

I INTRODUÇÃO

Em corporações modernas, é comum adotar uma estrutura que separa a propriedade da

administração. Embora os acionistas contem com direito a voto, a administração diária é, na verdade, realizada por executivos profissionais indicados pelo conselho administrativo (cujos integrantes são eleitos pelos acionistas). Essa situação abre caminho para conflitos de interesse: as decisões tomadas por executivos nem sempre são as que os acionistas tomariam caso estivessem gerindo a empresa (JENSEN; MECKLING, 1976). O CEO, por exemplo, pode comprar uma passagem aérea na primeira classe, enquanto um acionista, talvez, preferisse comprar uma passagem mais barata.

Essa diferença de interesses geralmente faz a empresa incorrer em custos de agência. Alguns desses custos são os custos com fiscalização: os acionistas estão dispostos a investir no desenvolvimento de um sistema de fiscalização capaz de evitar abusos por parte da administração. As demonstrações financeiras, elaboradas de acordo com algumas normas de divulgação financeira (ex. IFRS), fazem parte desse sistema de fiscalização. Embora elas caracterizem um custo (que acaba sendo pago pelos acionistas), permitem o acesso dos investidores às informações sobre a qualidade da administração da empresa. A fim de assegurar aos investidores que as normas de divulgação financeira adequadas foram realmente seguidas, a empresa geralmente contrata um auditor independente. Na verdade, os usuários das demonstrações financeiras dependem dos auditores para prevenirem e identificarem fraudes ou erros nas demonstrações (BEST; BUCKBY; TAN, 2001; LEE; GLOECK; PALANIAPPAN, 2007), embora “seja responsabilidade da administração desenvolver e implantar programas e controles para prevenir, deter e identificar fraudes.” (AICPA, 2002, p. 169).

Mesmo com os complexos sistemas de fiscalização adotados nas empresas de capital aberto, a fraude nas demonstrações financeiras (FDF) ainda assombra os usuários do documento. Alguns casos importantes de fraude contábil tornaram-se conhecidos no início da primeira década do século 21, quebrando a confiança dos investidores, à medida que os custos das fraudes são, em última instância, pagos pelos acionistas quando o preço das ações despensa após o anúncio das acusações

de fraude por parte da SEC (GERETY; LEHN, 1997). Algumas grandes empresas atualmente associadas à fraude nos Estados Unidos são Adelfia, Enron, Tyco e WorldCom. O surgimento das fraudes levou a uma reforma administrativa conhecida como Lei Sarbanes-Oxley (SOX), que estabeleceu novas normas de transparência para as empresas de capital aberto. Apesar dessas medidas, as fraudes ainda ocorrem (embora em número aparentemente menor depois de 2002) e, portanto, são ainda tema relevante para a pesquisa acadêmica.

A não identificação ou prevenção da fraude nas demonstrações financeiras pode prejudicar a reputação e a credibilidade da profissão do auditor (CHUI; PIKE, 2011). De fato, alguns argumentam que a fraude é o tendão de Aquiles da profissão do auditor, ao menos em parte, visto que a auditoria com base nos riscos é previsível (JAMAL, 2008). O uso de casos passados de fraude para construir modelos para estimar o risco de fraude não ajuda os auditores a se tornarem imprevisíveis, mas pode ajudá-los a entender as causas da FDF. A análise de regressão (isto é, modelos logit e probit) permite que se testem hipóteses sobre a influência de certas variáveis na ocorrência da fraude.

Um dos problemas dessa abordagem é a existência de casos de fraude ainda não descobertos, alguns dos quais, provavelmente, nunca serão. É possível ainda que a SEC opte por se concentrar em grandes casos de fraude, sendo, portanto, alguns dos casos de certa forma identificados, mas não totalmente investigados e denunciados nos Comunicados de Execução Contábil e de Auditoria (AAERs). Por fim, é possível ainda que a comissão suspeite de que algumas demonstrações financeiras estejam intencionalmente erradas, mas não consiga apresentar evidências que sirvam de apoio à suspeita. Não seria de surpreender se apenas uma pequena parte das fraudes contábeis fosse descoberta, totalmente investigada e divulgada ao público. Em uma análise de regressão, isso pode gerar problemas de erro de classificação na variável dependente. Há empresas indevidamente excluídas da relação de casos de fraude. Esses erros prejudicam o modelo e, dessa forma, as estimativas

dos parâmetros. É possível, portanto, que os testes das hipóteses sejam afetados, fazendo com que as conclusões possam enganar. Hausman, Abrevaya e Scott-Morton (1998) desenvolveram um método capaz de minimizar esse problema, além de estimar a probabilidade de um caso de fraude não ser descoberto. Este estudo examina os resultados desse método aplicado à FDF, a fim de avaliar se esses resultados são diferentes dos obtidos com um modelo logit não corrigido.

Este artigo conta com três objetivos. Nosso primeiro objetivo é estimar o percentual de fraudes em empresas dos Estados Unidos não divulgadas. Ao fazê-lo, podemos analisar a eficácia da SEC na identificação e repressão da FDF. Nosso segundo objetivo é verificar se os resultados obtidos por meio de um modelo logit tradicional são qualitativamente diferentes dos obtidos por meio do modelo logit com erro de classificação de parâmetros. Nosso último objetivo é a obtenção de testes de hipóteses confiáveis em relação a algumas variáveis que podem estar relacionadas à FDF.

2 REFERENCIAL TEÓRICO E VARIÁVEIS

A fraude, em um sentido amplo, é definida como a distorção consciente da verdade ou ocultação de fato relevante com o objetivo de induzir outras pessoas a agirem em detrimento dos próprios interesses (PEDNEAULT, 2009). Quando um executivo distorce intencionalmente as demonstrações financeiras, ele também comete fraude. Há uma ação consciente que induz outras pessoas a agirem em detrimento dos próprios interesses, como pagar um preço alto por ações ordinárias que, na verdade, valem menos. Esses casos são conhecidos como fraudes das demonstrações financeiras (ALBRECHT, W. S. et al., 2009).

É importante acrescentar que as FDF são cometidas pela violação dos Princípios Contábeis Geralmente Aceitos (GAAP). Caso não haja violação dos GAAP, é então um caso de gestão de rendimentos, que também é prejudicial aos investidores, mas recorre apenas a meios legais (DECHOW; SKINNER, 2000).

A fim de estimar o modelo logit para a detecção de fraude, é necessário identificar um conjunto de variáveis que expliquem a ocorrência da fraude. Existem conceitos e teorias que explicam o crime do colarinho branco, alguns dos quais podem ser empregados como ponto de partida para o entendimento das causas da fraude contábil, servindo como orientação para a seleção e organização de variáveis independentes. É comum usar o triângulo da fraude, originalmente proposto por Donald R. Cressey, para explicar a ocorrência do crime do colarinho branco. O autor acredita que esses crimes ocorram quando da presença de três fatores: pressão (necessidade financeira), oportunidade e racionalização. No início da década de 1980, Steve Albrecht adaptou a teoria para estudar as FDF (CHOO; TAN, 2007). Classificamos as variáveis utilizadas neste estudo nesses três grupos.

Deve-se observar que, ao estimar um modelo logit, uma observação é excluída caso haja algum valor faltando para qualquer variável. A fim de evitar o trabalho com uma amostra pequena, excluímos todas as variáveis com menos de 20 mil observações. Essas variáveis foram coletadas, mas não foram utilizadas nem mencionadas neste artigo.

2.1 Pressão

A pressão (necessidade financeira) está relacionada à situação da empresa e de seus administradores. Espera-se que as empresas que não tenham conseguido atingir as expectativas do mercado sejam mais suscetíveis à fraude, uma vez que os administradores estão sob pressão para que o desempenho seja bom. Por esse motivo, os indicadores de desempenho e a segurança financeira podem ser usados na detecção de fraude, além de outras evidências. A Enron, por exemplo, foi uma empresa de grande alavancagem que exigia que seus executivos obtivessem (e divulgassem) altos lucros a fim de pagar as participações e cumprir obrigações contratuais (CHOO; TAN, 2007).

Para capturar os elementos da pressão nesse modelo, usamos diversos indicadores diferentes. Um deles é modelo Z de Altman (**zscore**), um proxy para o risco de dificuldades

financeiras que costuma ser usado na pesquisa de FDF (ALTMAN, 1968). A expectativa é que as empresas com escore Z mais alto se encontrem em uma posição financeira melhor e, portanto, apresentem menor probabilidade de lançar mão da fraude, embora pesquisas anteriores mostrem que esse fato esteja relacionado negativamente à probabilidade de fraude em empresas gregas (KIRKOS; SPATHIS; MANOLOPOULOS, 2007; SPATHIS, 2002), mas estatisticamente não relacionado à fraude em empresas dos Estados Unidos (ERICKSON; HANLON; MAYDEW, 2006; FANNING; COGGER, 1998; SUMMERS; SWEENEY, 1998). Usamos ainda diversas outras variáveis que tentam avaliar a capacidade das empresas de cobrirem seus passivos, visto que empresas com dificuldade para pagar as dívidas (e, portanto, que têm problemas financeiros) podem recorrer à FDF para obterem recursos financeiros, tanto em mercados de capitais quanto por meio de empréstimos. Uma dessas variáveis é o caixa em relação ao total de ativos (**CASHTA**), proposta por Gaganis (2009), que descobriu haver uma relação negativa entre a variável e a fraude. Adotamos ainda a variação no caixa em relação ao total de ativos (Δ CASHTA). Outra variável relacionada à liquidez é o ativo circulante dividido pelo passivo circulante (**CACL**), também conhecida como razão circulante. Um **CACL** abaixo de 1 sugere que a empresa não possui ativos suficientes prontamente disponíveis para pagar dívidas de curto prazo e que, portanto, dá sinais de possíveis problemas de insolvência. A expectativa dessa variável e suas alterações (Δ CACL) é a de estar negativamente relacionada à fraude, embora haja estudos prévios indicando que a variável não se relacione estatisticamente com a FDF (BENEISH, 1999a; KIRKOS; SPATHIS; MANOLOPOULOS, 2007). Outra variável é o capital de giro (isto é, ativo circulante menos passivo circulante) dividido pelo total de ativos (**WCCTA**), usada por Beneish (1997, 1999a, 1999b), Fanning e Cogger (1998), Spathis (2002), Kirkos, Spathis e Manolopoulos (2007) e Gaganis (2009). Esperamos que a variável e sua variação (Δ WCCTA) esteja relacionada negativamente à fraude.

Fanning e Cogger (1998) e Kaminski, Wetzel e Guan (2004) usaram a razão formada pelas contas a receber divididas pelo total de ativos (**ARTA**). Acreditamos que as empresas com **ARTA** alta sejam financeiramente menos seguras e que, portanto, apresentem maior probabilidade de fraude. Uma variável que costuma ser usada na avaliação da segurança financeira é a **TLTA**, obtida pelo total do passivo em relação ao total do ativo (BENEISH, 1997, 1999a, 1999b; CRUTCHLEY; JENSEN, M. R. H.; MARSHALL, 2007; ERICKSON; HANLON; MAYDEW, 2006; ETTREDGE et al., 2008; JOHNSON; RYAN; TIAN, 2008; WANG, 2011). Usamos ainda a variação dessa variável ($\Delta TLTA$). Medidas bastante similares de alavancagem, usadas por Fanning e Cogger (1998), podem ser obtidas pela substituição do total de ativos pelo total do patrimônio líquido (**TLTE** e $\Delta TLTE$). Outro indicador de segurança financeira é a razão de vendas com as contas a receber (**SALAR**), usado por Summers e Sweeney (1998), Spathis (2002), Kirkos, Spathis e Manolopoulos (2007) e Skousen e Wright (2008). Esperamos que ele se relacione negativamente à probabilidade de fraude. Kaminski, Wetzel e Guan (2004) utilizaram ainda a razão para a segurança financeira, dada pela divisão das despesas de juros pelo total de passivos (**IETL**).

Outra variável que pode ser usada é a **FATA**, a razão de ativos fixos sobre o total de ativos (KIRKOS; SPATHIS; MANOLOPOULOS, 2007). É o ativo fixo (ex. equipamentos e fábricas) que geram receita para a empresa e, portanto, esperamos que empresas com valores mais altos investidos em ativo imobilizado tenham menor probabilidade de recorrer à fraude, uma vez que se encontram em posição melhor para a geração de receita no futuro. Segundo Kirkos, Spathis e Manolopoulos, a **FATA** não se relaciona com a fraude em empresas gregas. Usamos ainda a variação dessa variável ($\Delta FATA$). Outra variável, proposta por Gaganis (2009) e negativamente relacionada à **FATA**, é obtida pela divisão do ativo circulante pelo total de ativos (**CATA**). Sua variação também foi usada neste estudo ($\Delta CATA$). Gaganis descobriu que a **CATA** é significativamente menor em empresas envolvidas em FDF.

Alguns autores sugerem o uso do custo de mercadorias vendidas dividido pelas vendas (KAMINSKI; WETZEL; GUAN, 2004). Esperamos que essa variável (**COGSAL**) e sua variação ($\Delta COGSAL$) estejam positivamente relacionadas à probabilidade de fraude, visto que o aumento no custo de mercadorias vendidas pode sinalizar uma redução na capacidade da empresa de competir.

Spathis (2002) e Kirkos, Spathis e Manolopoulos (2007) usaram, em seus modelos, uma variável obtida pela divisão do lucro líquido pelas vendas. Usamos o EBIT (lucro antes dos juros e impostos) dividido pelas vendas (**EBITSAL**) como tentativa de prever a FDF. A explicação por trás dessa variável é que a redução da lucratividade cria uma pressão que pode levar a empresa a recorrer à fraude. Uma variável semelhante é a EBIT em relação ao total de ativos (**EBITTA**). Usamos ainda a variação dessa variável ($\Delta EBITTA$). Outra avaliação da lucratividade é o lucro líquido em relação ao total de ativos (**ROA**) e sua variação (ΔROA). Ao construir variáveis como a **EBITTA** e a **ROA**, devemos observar que as variáveis obtidas a partir da demonstração do resultado (como o lucro líquido e o EBIT) são atribuídas a todo o período, enquanto as variáveis obtidas com base no balanço patrimonial (como o total de ativos) referem-se a um momento específico. Para que o denominador possa ser comparado ao numerador, usamos a média do total de ativos no período pertinente à geração das variáveis **EBITTA** e **ROA**, assim como suas variações.

A literatura existente apresenta ainda variáveis com base no valor dos estoques. A presença de estoque em excesso pode sinalizar que a empresa não consegue sustentar o nível de previsão de vendas ou, simplesmente, é ineficiente. Fanning e Cogger (1998) e Gaganis (2009), usando o estoque dividido pelas vendas (**INVSAL**), descobriram evidências que sustentam essa hipótese. Por outro lado, em empresas gregas, essa variável parece ser estatisticamente insignificante (KIRKOS; SPATHIS; MANOLOPOULOS, 2007; SPATHIS, 2002). Summers e Sweeney (1998) argumentam que a FDF está relacionada ao crescimento do estoque ($\Delta INVSAL$). Usamos essa variável porque alguns setores podem apresentar

níveis de estoque mais elevados. Por fim, alguns autores usaram o estoque dividido pelo total de ativos (**INVTA**), mas descobriram que essa variável não era estatisticamente significativa (FANNING; COGGER, 1998; GAGANIS, 2009; KIRKOS; SPATHIS; MANOLOPOULOS, 2007).

2.2 Oportunidade

O segundo elemento do triângulo da fraude é a oportunidade, que não deve ser entendida somente como a falta de regulamentação ou má governança corporativa. Segundo Choo e Tan (2007), a grande ênfase no sucesso monetário por parte da sociedade americana pode levar os gestores a buscarem, de forma ativa, meios para contornar o mecanismo institucional criado para reprimir a fraude. Isso significa que, mesmo quando a oportunidade não está presente, em alguns casos, os gestores poderão criá-la. A pressão pelo sucesso gera oportunidades para o comportamento criminoso. Os gestores da Cendant, empresa acusada de fraude, mantinham uma planilha anual com as “oportunidades” de fraude disponíveis para inflar a receita operacional e relacionaram os valores que deveriam ser obtidos em cada uma dessas oportunidades.

Em nosso modelo, a variável que representa a oportunidade de fraude é um dummy para mudança do auditor (ETTREDGE et al., 2008; FANNING; COGGER, 1998; SKOUSEN; WRIGHT, 2008), aqui abreviada como **AUDCH**. Segundo Summers e Sweeney (1998), é possível que uma empresa que se envolva em fraude mude de auditor para diminuir a probabilidade de ser pega.

2.3 Racionalização

Os gestores envolvidos em fraudes (e os criminosos do colarinho branco em geral) “ajustam’ a construção simbólica de seu comportamento a fim de adaptar-se às expectativas sociais em geral,” em um processo comumente chamado de racionalização (COLEMAN, 1987, p. 410). Eles podem, por exemplo, afirmar que suas ações foram necessárias para que a empresa sobrevivesse e evitasse demissões. Ainda, em setores com

concorrência acirrada, os gestores podem argumentar que atos fraudulentos estão por toda a parte, interpretando-os como algo natural e esperado.

Para Skousen e Wright (2008), o nível de acumulações representa a forma de a administração tomar decisões sobre as demonstrações financeiras. Quando a gestão de rendimentos é vista como normal, ações mais sérias podem ser o próximo passo. Como o uso excessivo de acumulações costuma motivar relatórios de auditoria com modificação (FRANCIS; KRISHNAN, 1999), usamos uma variável dummy (**UQUAL**), que é igual a 1 se o relatório de auditoria apresentar um parecer sem ressalvas. O parecer é sem ressalvas quando as demonstrações financeiras não refletem restrições sem solução e o auditor não conta com exceções significativas. Os pareceres sem ressalvas com redação adicional recebem o código 0, visto que podem refletir o uso excessivo de acumulações. Em todos os outros casos (ex. opinião com ressalva ou adversa), a variável é igual a 0.

2.4 Estudos semelhantes

Há diversos artigos que utilizaram os modelos probit e logit para estimar a probabilidade de FDF ou identificar as variáveis que influenciam tal probabilidade. Alguns deles utilizaram ainda os AAERs para criar uma variável dependente (ABBOTT; PARK; PARKER, 2000; BRAZEL; JONES; ZIMBELMAN, 2009; CRUTCHLEY; JENSEN, M. R. H.; MARSHALL, 2007; ERICKSON; HANLON; MAYDEW, 2006; ETTREDGE et al., 2008; JOHNSON; RYAN; TIAN, 2008; LENNOX; PITTMAN, 2010; MILLER, 2006; SKOUSEN; WRIGHT, 2008; PERSONS, 1995). Outros autores usaram os casos de fraude divulgados pela imprensa, sozinhos ou juntamente com os AAERs (BEASLEY, 1996; BENEISH, 1997; LEE; INGRAM; HOWARD, 1999; SUMMERS; SWEENEY, 1998).

Os artigos que usam os modelos probit e logit costumam ter como objetivo verificar o impacto de uma determinada variável sobre a ocorrência da fraude. Erickson, Hanlon e Maydew (2006), por exemplo, estudaram o efeito das opções de compra de ações sobre a fraude. Os

resultados deste artigo indicaram que, ao contrário do que se costuma acreditar, não há evidências consistentes de que incentivos baseados em ações contribuem para a ocorrência da fraude contábil. Esses resultados foram confirmados pelos trabalhos de Gerety e Lehn (1997) e Crutchley, Jensen e Marshall (2007), mas não por Erickson, Hanlon e Maydew (2004). Mais recentemente, Feng et al. (2011) descobriram evidências que sustentam a hipótese de que os CFOs se envolvem em fraudes por sucumbir à pressão dos CEOs, e não por quererem aumentar sua renda a partir dos incentivos de títulos.

Há diversos trabalhos que empregaram metodologias para estimar a probabilidade que um caso de fraude não seja identificado (WANG, 2011). O autor usou o método proposto por Lin (1980), no qual a probabilidade de fraude e a probabilidade de identificação da fraude são estimadas separadamente. O modelo usado por Wang requer a identificação de um conjunto de variáveis que influencie a probabilidade de identificação de fraude. Nosso artigo, por outro lado, estimou a probabilidade incondicional de identificação de fraude (sem o uso de um conjunto adicional de variáveis independentes).

Alguns trabalhos usaram técnicas de inteligência artificial para identificar as fraudes. É comum o uso de redes neurais para esse fim (CHOI; GREEN, 1997; FANNING; COGGER, 1998; OGUT et al., 2009). Alguns estudos empregaram a lógica nebulosa (DESHMUKH; TALLURU, 1998), enquanto outros adotaram métodos híbridos como as redes neurais nebulosas (LIN; HWANG; BECKER, 2003). Por fim, há artigos que usaram diversos métodos de detecção de fraude diferentes com o intuito de compará-los (KIRKOS; SPATHIS; MANOLOPOULOS, 2007; GAGANIS, 2009).

Um dos problemas com os estudos que usam a inteligência artificial é a dificuldade de testar hipóteses. Não é possível, por exemplo, testar se a probabilidade de fraude está relacionada à remuneração do executivo. Por outro lado, diversos estudos concluíram que os métodos de inteligência artificial podem identificar a ocorrência de fraude com maior precisão quando

comparados aos modelos probit e logit (LIN; HWANG; BECKER, 2003; OGUT et al., 2009). A escolha de ferramentas depende dos objetivos do pesquisador. Neste artigo, como nosso objetivo é estimar a probabilidade de uma fraude passar sem ser detectada e testar a hipótese de que essa probabilidade seja diferente de zero, escolhemos o modelo logit.

3 MÉTODO

O presente artigo apresenta um modelo empírico. Os dados relacionados às demonstrações financeiras foram obtidos no Compustat. As próximas seções descrevem como os AAERs publicados pela SEC foram usados na identificação das empresas acusadas de fraude nos Estados Unidos e explicar o novo método usado para estimar a probabilidade de uma fraude não ser detectada.

3.1 Dados

As empresas acusadas de FDF foram identificadas por meio de uma análise dos AAERs publicados pela SEC de 1º de janeiro de 1988 a 23 de novembro de 2010. Todos os relatórios foram lidos em busca de distorções materialmente relevantes.

Se a empresa foi acusada de fraude, lemos cuidadosamente o comunicado a fim de identificar o primeiro ano da fraude, isto é, o primeiro ano em que uma distorção materialmente relevante foi publicada. Tomamos cuidado para verificar os exercícios fiscais das empresas acusadas. O Compustat adota uma regra simples: Caso o exercício fiscal se encerre entre 1º de janeiro e 31 de maio, é considerado o ano-calendário anterior. Uma demonstração encerrada em 31 de março de 2000, por exemplo, é armazenada na base de dados como sendo do ano 1999. Portanto, se a fraude tiver sido identificada pela primeira vez nessa demonstração, o primeiro ano da fraude será 1999. Em outras palavras, tomamos cuidado para relacionar a lista de fraudes e os exercícios com os dados fornecidos pelo Compustat.

Removemos todas as ocorrências de violação da Lei de Práticas de Corrupção Estrangeira (FCPA) da amostra. A FCPA proíbe a prática de suborno de oficiais estrangeiros. Como o suborno é ilegal, o dinheiro usado para esse fim não passa pela contabilização oficial; a empresa pode recorrer a algum tipo de violação de relatório financeiro a fim de esconder o destino desses recursos. Isso pode afetar de forma significativa as demonstrações financeiras da empresa. Infelizmente, nem todos os AAERs emitidos por violações da FCPA informam que demonstrações financeiras foram afetadas. Às vezes, o valor destinado ao suborno pode ser negligenciado. Por causa desses problemas, foram todos removidos da amostra.

Não levamos em consideração ainda o backdating de opções de compra de ações. O principal motivo foi a dificuldade de descobrir variáveis que possam ser explicadas relacionadas a esse tipo de fraude, o que é desenvolvido muito mais a fim de beneficiar alguns executivos do que enganar o público geral sobre a saúde financeira da empresa. A inclusão do backdating de opções de compra de ações neste estudo apenas misturariam dois fenômenos na mesma variável dependente. Entretanto, a prática do backdating de opções de compra de ações pode distorcer as demonstrações financeiras.

Devido a limitações técnicas, mantivemos na amostra apenas empresas com registro em três bolsas de valores (NYSE, NASDAQ e AMEX).

Isso significa que as empresas que negociadas no mercado de balcão foram excluídas, assim como aquelas sem dados disponíveis no Compustat.

Uma FDF costuma ser descoberta somente anos depois de começar. É difícil, portanto, obter uma amostra significativa de empresas que tenham emitido demonstrações financeiras falsificadas em exercícios recentes. Qualquer tentativa de fazê-lo aumenta o número de casos de erros de classificação, visto que muito mais empresa seriam inadequadamente classificadas como honestas. Assim, optamos também por remover da amostra os casos de fraude ocorridos antes de 1998 e depois de 2002. Como podemos verificar na Figura 1, o número de fraudes descobertas depois de 2002 aparentemente caiu de forma abrupta. Em vez de sinalizar uma verdadeira redução do número de fraudes, essa queda mais provavelmente mostra a incapacidade da SEC de denunciar casos recentes.

As empresas financeiras – as que possuem códigos de Classificação Industrial Padrão (SIC) de 6000 a 6999 – também foram removidas da amostra. A amostra final contém 118 empresas que atendem a todos os requisitos.

Os modelos não foram estimados somente com as empresas acusadas de fraude. A população deste estudo consiste em todas as empresas com demonstrações financeiras disponíveis no Compustat no período de 1998 a 2002. As empresas com dados de apenas alguns desses cinco anos foram incluídas.

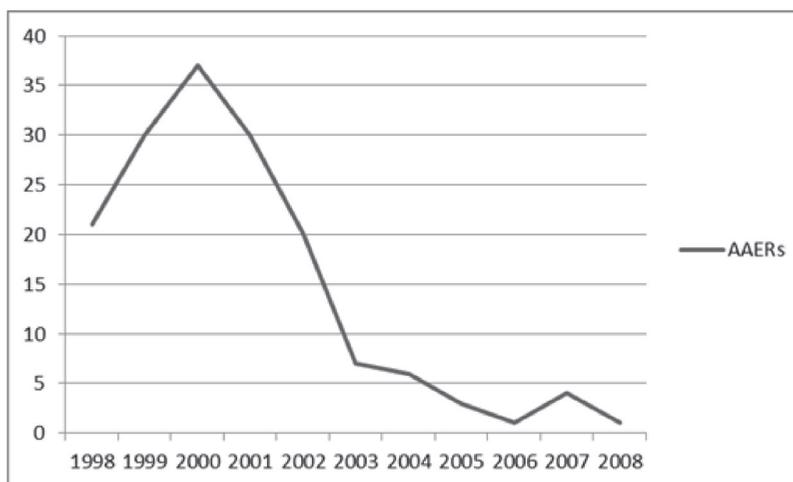


FIGURA 1 – AAERs emitidos entre 1998 e 2008

Fonte: Dos autores.

3.2 Análise fatorial

A análise fatorial é uma técnica multivariada que sintetiza a relação observada em um conjunto de variáveis inter-relacionadas, permitindo a identificação de fatores comuns. Segundo Jöreskog (1969), em uma análise fatorial com p variáveis e k fatores, o modelo básico é:

$$\mathbf{z}_i = \Gamma \mathbf{f}_i + \boldsymbol{\varepsilon}_i \quad (1)$$

em que \mathbf{z}_i é um vetor de ordem p de variáveis observadas para observação i , \mathbf{f}_i é um vetor de ordem $k < p$ de escores fatoriais latentes comuns para a observação i , e $\boldsymbol{\varepsilon}_i$ é um vetor de ordem p de escores únicos para observação i . Esses escores únicos podem ser vistos como “termos de erro”: Representam a porção de \mathbf{z}_i que não pode ser explicada por \mathbf{f}_i . A letra maiúscula grega gama (Γ) é uma matriz $p \times k$ da carga fatorial.

Como sugere o modelo, a análise fatorial assume que é possível representar um conjunto de p variáveis usando um número menor de fatores intrínsecos k . Contudo, para que a análise fatorial seja bem-sucedida, deve haver correlações razoavelmente fortes entre as variáveis usadas na análise. Em outras palavras, a matriz de correlação dessas variáveis deve ser diferente da matriz identidade. Para verificar essa premissa, usamos o teste de esfericidade de Bartlett, que rejeitou as hipóteses nulas de esfericidade ($p = ,000$). Concluímos que a análise fatorial é adequada ao nosso conjunto de dados.

Há diversos procedimentos de extração de fatores que podem ser utilizados. O método de máxima verossimilhança (ML) tem a vantagem de permitir o teste da materialidade estatística da carga fatorial. Esse método, no entanto, exige a premissa da normalidade multivariada (FABRIGAR et al., 1999). Por meio do teste Doornik-Hansen (DOORNIK; HANSEN, 2008), descobrimos que as variáveis usadas neste artigo rejeitaram a hipótese nula da normalidade multivariada ($p = ,000$). Para evitar resultados distorcidos, optamos pelo uso do método de fatores principais (PF) para a extração de fatores. Esse método não tem premissas distribucionais.

3.3 Estimativa do modelo

Nos modelos probit e logit, a variável dependente \tilde{y}_i costuma ser expressa como uma função de variável latente e observável I_i e um limiar desconhecido I_i^* (GUJARATI, 1995; POWERS; XIE, 2000):

$$\begin{aligned} \tilde{y}_i &= 1 \text{ se } I_i > I_i^* \\ \tilde{y}_i &= 0 \text{ se } I_i < I_i^* \end{aligned} \quad (2)$$

Quando há erro de classificação na variável dependente, as condições acima não são respeitadas. A variável observada y_i não é a mesma que a variável verdadeira \tilde{y}_i . O modelo agora tem dois novos parâmetros α_0 e α_1 :

$$\begin{aligned} \alpha_0 &= \Pr(y_{it} = 1 | \tilde{y}_{it} = 0) \\ \alpha_1 &= \Pr(y_{it} = 0 | \tilde{y}_{it} = 1) \end{aligned} \quad (3)$$

Em estudos sobre a FDE, o erro mais importante é o demonstrado como parâmetro α_1 na segunda equação. Esse erro é a probabilidade de a empresa que tenha emitido demonstrações financeiras falsas não ser classificada com tal (erro tipo II). A expectativa é de que o outro erro (tipo I) seja bastante raro e próximo a zero. Portanto, para facilitar a convergência do modelo, este artigo adota a premissa de que $\alpha_0 = 0$. Sem essa restrição, o valor esperado de y_i em um modelo logit é (HAUSMAN; ABREVAYA; SCOTT-MORTON, 1998):

$$E(y_i | \mathbf{x}_i) = \Pr(y_i = 1 | \mathbf{x}_i) = \alpha_0 + (1 - \alpha_0 - \alpha_1) \Lambda(\mathbf{x}_i' \boldsymbol{\beta}) \quad (4)$$

A função de log-verossimilhança foi efetivamente maximizada neste artigo, com base na equação 4, mas com a restrição $\alpha_0 = 0$, é:

$$\ln L = \sum_{y_i=0} \ln[1 - (1 - \alpha_1) \Lambda(\mathbf{x}' \boldsymbol{\beta})] + \sum_{y_i=1} \ln[(1 - \alpha_1) \Lambda(\mathbf{x}' \boldsymbol{\beta})] \quad (5)$$

A função 5 foi maximizada por meio do método Newton-Raphson.

4 RESULTADOS

Extraímos 7 fatores das 28 variáveis usando o método de fatores principais. Esses 7 fatores são responsáveis por 89% da variação das variáveis originais (Tabela 1). Isso significa que não há grande perda de informações quando substituímos as 28 variáveis originais por esses 7 fatores no modelo logit.

TABELA 1 – Variação explicada pela análise fatorial (método de fatores principais)

Fator	Autovalor	Proporção	Cumulativo
Fator 1	3,39808	,2197	,2197
Fator 2	2,79139	,1804	,4001
Fator 3	2,22932	,1441	,5442
Fator 4	1,63002	,1054	,6496
Fator 5	1,41903	,0917	,7413
Fator 6	1,28765	,0832	,8245
Fator 7	1,06663	,0689	,8935

Fonte: Dos autores.

O primeiro fator (capaz de capturar quase 22 da variação das variáveis originais) parece estar relacionado positivamente com a saúde financeira das empresas. Relaciona-se positivamente com as variáveis **EBITTA**, **ROA**, **zscore**, **WCTA**, Δ WCTA. Além disso, relaciona-se negativamente com as variáveis **TLTA** e Δ TLTA. As empresas com menos dívida e **ROA** alto costumam apresentar valores do Fator 1 mais altos.

O Fator 2 captura 18% da variação das variáveis utilizadas. Relaciona-se positivamente com **COGSAL**, **CASHTA** e **INVSAL**, e negativamente com as variáveis Δ COGSAL e **EBITSAL**. Esses resultados mistos sugerem que as empresas com mais recursos investidos em alguns tipos de ativo circulante (caixa e estoque) e com custos de vendas decrescentes (mas ainda altos) tendem a apresentar valores mais altos para o Fator 2. Esse fator relaciona-se negativamente com a avaliação da lucratividade (**EBITSAL**).

O Fator 3 relaciona-se positivamente com a **CATA** e negativamente com a **FATA**. As empresas com mais ativos circulantes tendem a apresentar valores mais altos para o Fator 3. A variável Δ TLTE, embora bem capturada pelo Fator

A Tabela 2 mostra a carga fatorial de cada variável usada na análise fatorial (a matriz Γ na Eq. 1). Cada variável é igual à soma dos produtos dos fatores pela carga fatorial correspondente, mais seu escore único. Ao analisar os fatores absolutos mais altos de cada variável, podemos entender a composição de cada fator.

3, conta com carga fatorial muito baixa para ser considerada significativa.

O Fator 4 relaciona-se positivamente com Δ CASHTA, Δ CATA e **IETL**. Além disso, relaciona-se negativamente com Δ FATA. Esse fator captura aumentos de caixa e ativo circulante. Captura bem os aumentos e reduções das variáveis incluídas no Fator 3. As empresas com Fator 4 alto estão mudando a composição do ativo e incorrendo em mais despesas com juros.

O Fator 5 relaciona-se positivamente com a Δ INVSAL e negativamente com a Δ EBITTA e Δ ROA. É uma variável que captura reduções na lucratividade e aumentos no estoque. Esse fator captura aumentos e reduções em algumas das variáveis incluídas no Fator 1.

Não há variáveis que tenham carga fatorial absoluta mais alta no Fator 7. As cargas fatoriais mais altas associadas ao Fator 7 são Δ TLTA (0,406), **WCTA** (0,368) e Δ ROA (0,377). Parece que as empresas que aumentaram o passivo (mas ainda com ativo circulante suficiente para cobrir o passivo circulante) e melhoraram a lucratividade tendem a contar com um Fator 7 mais alto. As empresas com Fator 7 alto estão se tornando mais

lucrativas, possivelmente por causa da alavancagem financeira.

Realizamos uma regressão logit stepwise, começando com todos os fatores e três variáveis dummies que não foram incluídas na análise fatorial. A variável menos significativa (com valor p mais elevado) foi eliminada e o modelo foi estimado novamente. Esse procedimento de eliminação foi repetido até que todas as variáveis restantes tivessem valor p abaixo de 10%. O

modelo resultante encontra-se na Tabela 3. Deve-se observar que a amostra em cada modelo intermediário (e no final também.) continha apenas as observações com dados disponíveis para todas as variáveis usadas no modelo inicial. O acréscimo de observações depois de excluir uma variável, embora possível, poderia gerar estimativas não confiáveis. Dado o grande número de observações faltantes, a amostra continha 41 dos 118 casos de fraude originalmente identificados.

TABELA 2 – Carga fatorial e variações únicas

Variável	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6	Fator 7	Unicidade
COGSAL _{t-1}	-0,434	0,673	-0,504	-0,121	0,156	0,019	0,024	0,065
Δ COGSAL _{t-1}	0,273	-0,460	0,420	0,139	-0,164	-0,088	-0,075	0,479
EBITSAL _{t-1}	0,439	-0,676	0,482	0,114	-0,154	0,011	0,015	0,080
EBITTA _{t-1}	0,572	-0,249	-0,127	-0,182	0,271	0,023	0,283	0,408
Δ EBITTA _{t-1}	-0,218	0,172	0,006	0,223	-0,466	0,235	0,261	0,533
ROA _{t-1}	0,726	0,188	-0,261	-0,187	0,329	0,018	0,021	0,226
Δ ROA _{t-1}	-0,447	0,132	0,191	0,255	-0,468	0,180	0,377	0,287
zscore _{t-1}	0,299	0,145	-0,080	0,003	-0,179	-0,203	-0,002	0,810
CACL _{t-1}	0,181	0,324	0,129	0,047	-0,093	-0,356	-0,104	0,698
Δ CACL _{t-1}	-0,050	-0,087	-0,018	-0,034	0,022	0,159	0,041	0,961
CASHTA _{t-1}	0,175	0,556	0,299	0,186	-0,077	-0,479	-0,284	0,220
Δ CASHTA _{t-1}	0,197	0,269	0,175	0,541	0,345	-0,017	0,211	0,403
WCCTA _{t-1}	0,660	0,310	-0,169	-0,156	-0,328	-0,153	0,368	0,148
Δ WCCTA _{t-1}	0,516	0,197	-0,281	0,404	-0,182	0,429	-0,275	0,160
TLTA _{t-1}	-0,680	-0,210	0,317	0,149	0,303	0,184	-0,297	0,157
Δ TLTA _{t-1}	-0,561	-0,156	0,349	-0,290	0,147	-0,413	0,406	0,099
TLTE _{t-1}	0,001	-0,003	0,003	-0,009	0,003	0,012	0,004	1,000
Δ TLTE _{t-1}	-0,003	-0,008	-0,009	0,007	0,000	-0,005	0,000	1,000
ARTA _{t-1}	0,095	0,141	0,385	-0,379	0,069	0,461	0,058	0,459
INVSAL _{t-1}	-0,109	0,248	-0,163	-0,132	0,137	0,076	0,053	0,855
Δ INVSAL _{t-1}	-0,014	0,120	-0,137	-0,070	0,142	0,039	0,062	0,936
INVTAA _{t-1}	0,117	0,144	0,251	-0,293	0,055	0,328	0,151	0,683
CATA _{t-1}	0,252	0,618	0,635	-0,255	0,020	0,093	-0,110	0,066
Δ CATA _{t-1}	0,153	0,216	0,232	0,505	0,423	0,026	0,281	0,363
FATA _{t-1}	-0,178	-0,478	-0,488	0,250	0,009	-0,120	0,113	0,413
Δ FATA _{t-1}	-0,105	-0,181	-0,135	-0,364	-0,281	-0,059	-0,173	0,694
SALAR _{t-1}	-0,005	-0,018	-0,022	0,031	0,010	-0,014	0,019	0,998
IETL _{t-1}	-0,065	-0,030	-0,058	0,098	0,036	0,026	0,043	0,978

Nota. O fator com a carga fatorial absoluta mais elevada em cada linha encontra-se em negrito.

Fonte: Dos autores.

No modelo final, cinco fatores e uma variável dummy (**UQUAL**) foram significantes ao nível de 10%. O modelo como um todo foi também significativo, com o teste da razão de

verossimilhança rejeitando a hipótese nula ao nível de 1% ($p = ,0024$); isso significa que o modelo final enquadra os dados melhor que o modelo com apenas um intercepto.

TABELA 3 – Modelo logit sem parâmetro de erro de classificação

Variável	Razão de chances	Erro padrão	z	P> z
Fator 1	1,5426	0,3470	1,9300	,054
Fator 2	0,5671	0,1769	-1,8200	,069
Fator 4	0,3444	0,1166	-3,1500	,002
Fator 6	0,4695	0,1561	-2,2700	,023
Fator 7	3,7244	0,7088	2,8700	,004
UQUAL	0,3488	0,1135	-3,2400	,001

LR χ^2 : 20,39
 Prob > χ^2 : ,0024
 Log-verossimilhança: -259,47

Fonte: Dos autores.

Considerando a possibilidade de erro de classificação, usamos o modelo da Tabela 3 como modelo inicial e realizamos um procedimento stepwise. Esse procedimento excluiu três fatores (1, 2 e 6) e o modelo resultante encontra-se na Tabela 4. Os fatores restantes (4 e 7), embora ainda rejeitem a hipótese nula do teste z, apresentam valores p mais elevados no modelo sem um parâmetro de erro de classificação. Por outro lado, a variável **UQUAL** apresentou um valor p mais baixo. Os sinais das razões dessas variáveis permaneceram inalterados. Como o teste Wald rejeitou a hipótese nula, concluímos que ao menos um das razões das variáveis utilizadas é estatisticamente diferente de zero. O teste Wald é equivalente assintótico do teste LR usado no modelo sem parâmetro de erro de classificação (ENGLE, 1984).

As empresas com Fator 4 baixo e Fator 7 alto apresentam maior probabilidade de recorrerem

à fraude. *Ceteris paribus*, cada aumento de 1 no Fator 7 faz com que a empresa apresente probabilidade 2,6 vezes maior de se envolver em uma FDF. Como o Fator 4 e a **UQUAL** apresentam razões de chances abaixo de 1, decidimos aumentar as razões de chances à força -1 para elucidar melhor sua influência sobre a probabilidade de fraude. Dessa forma, cada redução de 1 no Fator 4 faz com que a empresa apresente probabilidade 2,49 vezes ($0,4014^{-1}$) maior de se envolver em uma fraude. Da mesma forma, a empresa que não recebeu um parecer com ressalva apresenta probabilidade 5,12 vezes ($0,1955^{-1}$) maior de se envolver em uma fraude.

O parâmetro de erro de classificação (α) estimado foi 98,57% e é estatisticamente diferente de zero ($p = ,000$). Isso significa que apenas uma pequena parte (1,43%) dos casos de fraude é descoberta e recebe a acusação da SEC.

TABELA 4 – Modelo Logit com parâmetro de erro de classificação

Variável dependente: AAER				
Observações: 10858				
Período: 1998-2002				
Variável	Razão de chances	Erro padrão	z	P> z
Fator 4	0,4014		-2,27	0,023
Fator 7	2,6047		1,69	0,091
UQUAL	0,1955		-3,97	0,000
Parâmetro de erro de classificação	Razão	Erro padrão	z	P> z
α_1	0,9855	0,0030	297,72	0,000

Wald χ^2 : 18,58
 Prob > χ^2 : ,0003
 Log-verossimilhança: -261,99

Fonte: Dos autores.

5 DISCUSSÃO E CONCLUSÃO

Nosso primeiro objetivo foi estimar o percentual de casos não detectados de fraudes em empresas dos Estados Unidos. O parâmetro de erro de classificação α_1 foi estimado em 98,57% e mostrou-se estatisticamente diferente de zero ($p = 0,000$). Esse resultado confirma o sugerido pelo senso comum: A SEC não consegue denunciar todos os casos de fraude por meio dos AAERs. As empresas envolvidas em FDF têm (*a priori*) uma probabilidade de 1,43% de serem pegas. O intervalo de confiança de 95% desse parâmetro vai de 97,90% a 99,22%.

Devemos observar que esses resultados não são uma afirmação sobre o que acontece dentro da comissão ou sua disposição para punir a fraude financeira. É possível que a SEC não tenha conhecimento de todos os casos de fraude não relatados, ou talvez tenha conhecimento de alguns, mas prefira se concentrar em casos maiores. De fato, deixar alguns (ou muitos) crimes sem punição pode ser aceitável do ponto de vista econômico, visto que há custos associados à investigação e à impetração de processos. Como Becker (1968) deixa claro em sua abordagem econômica ao crime, tentar punir todos os crimes, de assassinatos a pequenos furtos, é muito pouco provável do ponto de vista da sociedade. A mesma explicação pode ser aplicada ao crime do colarinho branco.

Nosso segundo objetivo foi verificar se o modelo logit com o parâmetro de erro de classificação produziria resultados diferentes em comparação ao modelo tradicional. Não realizamos testes a fim de verificar se as estimativas desses parâmetros são estatisticamente diferentes, mas os resultados são, na verdade, qualitativamente diferentes: três variáveis (fatores) foram excluídas no modelo de erro de classificação. Esses resultados sugerem que as conclusões obtidas por meio de modelos de detecção de fraude não corrigidos poderiam simplesmente não ser válidos.

Nosso último objetivo foi a obtenção de testes de hipóteses confiáveis em relação a algumas variáveis que podem estar relacionadas à FDF. Descobrimos três variáveis relacionadas à fraude ($p < ,10$). O Fator 4 relaciona-se negativamente com a fraude, indicando que as empresas com aumento do caixa em relação ao total de ativos (Δ CASHTA) apresentam menor probabilidade de se envolver em FDF. Como as empresas com mais caixa costumam estar sob menor estresse financeiro, esse resultado é compatível com nossas expectativas. Gaganis (2009), contudo, descobriu que a razão do caixa em relação ao total de ativos (CASHTA) relaciona-se negativamente com a fraude. Devemos observar, no entanto, que o autor não usou o aumento no caixa em relação ao total de ativos como variável e, portanto, os resultados não são diretamente comparáveis. O

mesmo cuidado se aplica a outras análises comparativas aqui apresentadas.

O Fator 4 indica ainda que as empresas com aumento do ativo circulante em relação ao total de ativos ($\Delta CATA$) apresentam menor probabilidade de se envolverem em FDF. Esperávamos resultados diferentes, visto que as empresas com imobilizado maior (e menos ativo circulante) apresentam maior capacidade de gerar receitas reais a partir de suas operações. Esses resultados são, entretanto, consistentes com Gaganis (2009), que descobriu que a **CATA** relaciona-se negativamente com a FDF.

O Fator 7 relaciona-se positivamente com a fraude, o que revela que as empresas com maior retorno sobre ativos (ΔROA) e alavancagem (avaliadas pelo total do passivo em relação ao total de ativos - $\Delta TLTA$) apresentam maior probabilidade de se envolverem em fraude. Embora o **ROA** seja amplamente usado na literatura, ainda não está claro como ele afeta a probabilidade de FDF. Segundo Summers e Sweeney (1998), o fator relaciona-se positivamente com a fraude, enquanto Crutchley, Jensen e Marshall (2007) e Gaganis (2009) sustentam que a relação é negativa. Esperávamos que o fator se relacionasse de forma negativa com a fraude, visto que uma redução da lucratividade deve pressionar os gestores para que eles se envolvam em FDF a fim de agradar os acionistas.

A maior parte dos estudos anteriores nos Estados Unidos não descobriu associação significativa entre a alavancagem (**TLTA**) e a FDF (CRUTCHLEY; JENSEN; MARSHALL, 2007; ERICKSON; HANLON; MAYDEW, 2006; ETTREDGE et al., 2008), embora Beneish (1999a) tenha descoberto uma associação positiva usando uma amostra mais antiga. Em mercados emergentes, como a Grécia e a China, a TLTA, na verdade, relaciona-se positivamente com a fraude (CHEN et al., 2006; KIRKOS; SPATHIS; MANOLOPOULOS, 2007; SPATHIS, 2002). Nossos resultados contribuem para a conclusão de que a alavancagem parece exercer pressão para o envolvimento em fraude apenas em mercados emergentes.

É interessante notar que, como primeira diferença (variação), as variações tiveram um papel

considerável na extração de ambos os fatores (4 e 7). Esses resultados podem sugerir que a fraude pode estar relacionada a variações de desempenho, especialmente reduções no passivo e melhoria da lucratividade dado o aumento da alavancagem. Isso pode ainda explicar por que nossos resultados costumam ser diferentes dos artigos já publicados, uma vez que eles raramente usaram as variações das variáveis. Sugerimos que estudos futuros se beneficiem com uma mudança para modelos que se concentrem nas mutações da posição financeira das empresas.

Os pareceres de auditoria também se relacionam à fraude, visto que um parecer adverso indica que as demonstrações financeiras apresentam distorções materialmente relevantes e um parecer com ressalvas indica que pode haver desvios em relação às normas contábeis adotadas. Em nosso estudo, usamos uma variável dummy (**UQUAL**) igual a um sempre que um parecer de auditoria não apresentava ressalva e zero de outra forma. Essa variável dummy foi usada por Skousen e Wright (2008) como um proxy para a racionalização, e a análise que fizeram dessa variável não rejeitou a hipótese nula do teste Wilcoxon ($z = -0,814$ e $p = ,208$). Repetimos o teste Wilcoxon e chegamos a resultados muito consistentes ($z = -1,30$, $p = ,1921$). Como Skousen e Wright adotaram um critério severo para a inclusão de uma variável em sua regressão logit ($p < ,15$ no teste Wilcoxon), eles não testaram se essa variável era relevante em um modelo multivariado. Descobrimos que, ao controlar o efeito de outras variáveis (incluídos nos fatores 4 e 7), o **UQUAL**, na verdade, relaciona-se negativamente com a probabilidade de FDF. Nossos resultados mostram que empresas com relatórios de auditoria limpos apresentam menor probabilidade de se envolverem em fraudes, conforme esperado. O resultado sugere ainda que o uso de testes univariados para decidir que variáveis devem ser usadas em um modelo multivariado pode levar à omissão de variáveis relevantes.

Recomendamos precaução no uso de modelos de escolha binária não corrigidos como ferramenta de apoio a decisões na identificação de fraude ou na avaliação do risco de fraude. É

mais adequado usar uma especificação que considere a possibilidade de problemas de erros de classificação e fazê-lo pode ter como resultado um modelo mais confiável. É importante ainda evitar as estimativas de probabilidade de fraude com subestimação de valores gerada por modelos tradicionais. Essas medidas podem contribuir para avaliações mais precisas sobre a probabilidade de fraude quando do uso de modelos econométricos. A adoção do modelo de erro de classificação é bastante direta, embora ainda não integrada ao software estatístico e econométrico, e poderia se tornar um procedimento padrão nas pesquisas sobre FDF.

É importante enfatizar que esse não é o único método que pode ser empregado para evitar os problemas resultantes do erro de classificação. Wang (2011) conseguiu estimar a probabilidade condicional de erro de classificação por meio de um método diferente. Vale ressaltar, contudo, que o método proposto por Wang requer a identificação de variáveis relacionadas à probabilidade de identificação de fraude. Embora isso possa permitir o teste de hipóteses de pesquisa interessantes, é uma responsabilidade desnecessária para os pesquisadores interessados apenas nas variáveis relacionadas à probabilidade de fraude.

À medida que as pesquisas avançam, será possível usar o método aqui adotado para estudar problemas relacionados à eficiência das agências de regulamentação na repressão da fraude. Testar variações no parâmetro de erro de classificação ao longo do tempo pode permitir conclusões sobre a eficiência de determinadas políticas relacionadas à fraude. Quando houver mais dados disponíveis, será possível testar, por exemplo, se a Lei Sarbanes-Oxley aumentou a capacidade da SEC de identificar e punir a FDF.

O método empregado aqui, entretanto, não é capaz de corrigir problemas de especificação do modelo. Se o modelo for subespecificado (isto é, houver variáveis faltantes), os resultados serão inconsistentes. Há ainda a necessidade de desenvolvimento teórico sobre as causas da fraude que nos permita identificar modelos melhores.

REFERÊNCIAS

- ABBOTT, L. J.; PARK, Y.; PARKER, S. The effects of audit committee activity and independence on corporate fraud. **Managerial Finance**, Bradford v. 26, n. 11, p. 55-68, 2000.
- ALBRECHT, W. S. et al. **Fraud examination**. 3rd. ed. Mason: South-Western Cengage, 2009.
- ALTMAN, E. I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. **The Journal of Finance**, Malden, v. 23, n. 4, p. 589-609, Sept. 1968.
- AMERICAN INSTITUTE OF CERTIFIED PUBLIC ACCOUNTANTS (AICPA). **Consideration of fraud in a financial statement audit**. AU Section 316. New York: AICPA, 2002.
- BEASLEY, M. S. An empirical analysis of the relation between the board of director composition and financial statement fraud. **The Accounting Review**, Sarasota, v. 71, n. 4, p. 443-465, Oct. 1996.
- BECKER, G. S. Crime and punishment: an economic approach. **Journal of Political Economy**, Chicago, v. 76, n. 2, p. 169-217, Mar./Apr. 1968.
- BENEISH, M. D. Detecting GAAP violation: implications for assessing earnings management among firms with extreme financial performance. **Journal of Accounting and Public Policy**, New York, v. 16, n. 3, p. 271-309, Autumn 1997.
- _____. The detection of earnings manipulation. **Financial Analysts Journal**, Charlottesville, v. 55, n. 5, p. 24-36, Sept./Oct. 1999a.
- _____. Incentives and penalties related to earnings overstatements that violate GAAP. **The Accounting Review**, Sarasota, v. 74, n. 4, p. 425-457, Oct. 1999b.
- BEST, P. J.; BUCKBY, S.; TAN, C. Evidence of the audit expectation gap in Singapore. **Managerial Auditing Journal**, Bradford, v. 16, n. 3, p. 134-144, 2001.

- BRAZEL, J. F.; JONES, K. L.; ZIMBELMAN, M. F. Using nonfinancial measures to assess fraud risk. **Journal of Accounting Research**, Chicago, v. 47, n. 5, p. 1135-1166, July 2009.
- CHEN, G.; FIRTH, M.; GAO, D.; RUI, O. Ownership structure, corporate governance, and fraud: evidence from China. **Journal of Corporate Finance**, London, v. 12, n. 3, p. 424-448, June 2006.
- CHOI, J. H.; GREEN, B. P. Assessing the risk of management fraud through neural network technology. **Auditing: A Journal of Practice & Theory**, Sarasota, v. 16, n. 1, p. 14, Mar. 1997.
- CHOO, F.; TAN, K. An "American dream" theory of corporate executive fraud. **Accounting Forum**, [S.l.], v. 31, n. 2, p. 203-215, June 2007.
- CHUI, L.; PIKE, B. Auditors' responsibility for fraud detection: new wine in old bottles? In: AMERICAN ACCOUNTING ASSOCIATION - ANNUAL MEETING, 2011, Denver. **Proceedings...** Sarasota: American Accounting Association – AAA, 2011.
- COLEMAN, J. W. Toward an integrated theory of white-collar crime. **American Journal of Sociology**, Chicago, v. 93, n. 2, p. 406-439, Sept. 1987.
- CRUTCHLEY, C. E.; JENSEN, M. R. H.; MARSHALL, B. B. Climate for scandal: corporate environments that contribute to accounting fraud. **The Financial Review**, Knoxville, v. 42, n. 1, p. 53-73, Feb. 2007.
- DECHOW, P. M.; SKINNER, D. J. Earnings management: reconciling the views of accounting academics, practitioners, and regulators. **Accounting Horizons**, Sarasota, v. 14, n. 2, p. 235-250, June 2000.
- DESHMUKH, A.; TALLURU, L. A rule-based fuzzy reasoning system for assessing the risk of management fraud. **Intelligent Systems in Accounting, Finance & Management**, Chichester, v. 7, n. 4, p. 223-241, Dec. 1998.
- DOORNIK, J. A.; HANSEN, H. An omnibus test for univariate and multivariate normality. **Oxford Bulletin of Economics and Statistics**, Oxford, v. 70, p. 927-939, Dec. 2008.
- ENGLE, R. F. Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. In: GRILICHES, Z.; INTRILIGATOR, M. D. (Eds). **Handbook of econometrics**. Amsterdam: North-Holland, 1984. v. 2, p. 775-826.
- ERICKSON, M.; HANLON, M.; MAYDEW, E. L. How much will firms pay for earnings that do not exist? Evidence of taxes paid on allegedly fraudulent earnings. **The Accounting Review**, Sarasota, v. 79, n. 2, p. 387-408, Apr. 2004.
- _____; _____. Is there a link between executive equity incentives and accounting fraud? **Journal of Accounting Research**, Chicago, v. 44, n. 1, p. 113-143, Mar. 2006.
- ETTREDGE, M. L. et al. Is earnings fraud associated with high deferred tax and/or book minus tax levels? **Auditing: A Journal of Practice & Theory**, Sarasota, v. 27, n. 1, p. 1-33, May 2008.
- FABRIGAR, L. R. et al. Evaluating the use of exploratory factor analysis in psychological research. **Psychological Methods**, Washington, D.C, v. 4, n. 3, p. 272-299, Sept. 1999.
- FANNING, K.; COGGER, K. O. Neural network detection of management fraud using published financial data. **Intelligent Systems in Accounting, Finance & Management**, Chichester, v. 7, n. 1, p. 21-41, Mar. 1998.
- FENG, M. et al. Why do CFOs become involved in material accounting manipulations? **Journal of Accounting and Economics**, Amsterdam, v. 51, n. 1-2, p. 21-36, Feb. 2011.
- FRANCIS, J. R.; KRISHNAN, J. Accounting accruals and auditor reporting conservatism. **Contemporary Accounting Research**, Toronto, v. 16, n. 1, p. 135-165, Mar. 1999.

GAGANIS, C. Classification techniques for the identification of falsified financial statements: a comparative analysis. **Intelligent Systems in Accounting, Finance & Management**, Chichester, v. 16, n. 3, p. 207-229, July 2009.

GERETY, M.; LEHN, K. The causes and consequences of accounting fraud. **Managerial and Decision Economics**, Chichester, v. 18, n. 7/8, p. 587-599, Nov./Dec. 1997.

GUJARATI, D. N. **Basic econometrics**. 3rd ed. New York: McGraw-Hill, 1995.

HAUSMAN, J. A.; ABREVAYA, J.; SCOTT-MORTON, F. M. Misclassification of the dependent variable in a discrete-response setting. **Journal of Econometrics**, Amsterdam, v. 87, n. 2, p. 239-269, Sept. 1998.

JAMAL, K. Mandatory audit of financial reporting: a failed strategy for dealing with fraud. **Accounting Perspectives**, Toronto, v. 7, n. 2, p. 97-110, May 2008.

JENSEN, M. C.; MECKLING, W. H. Theory of the firm: managerial behavior, agency costs and ownership structure. **Journal of Financial Economics**, Amsterdam, v. 3, n. 4, p. 305-360, Oct. 1976.

JOHNSON, S. A.; RYAN, H. E.; TIAN, Y. S. Managerial incentives and corporate fraud: the sources of incentives matter. **Review of Finance**, Dordrecht, Oxford, v. 13, n. 1, p. 115-145, May 2008.

JÖRESKOG, K. G. A general approach to confirmatory maximum likelihood factor analysis. **Psychometrika**, New York, v. 34, n. 2, p. 183-202, June 1969.

KAMINSKI, K. A.; WETZEL, T. S.; GUAN, L. Can financial ratios detect fraudulent financial reporting? **Managerial Auditing Journal**, Bradford, v. 19, n. 1, p. 15-28, 2004.

KIRKOS, E.; SPATHIS, C.; MANOLOPOULOS, Y. Data mining techniques for the detection of

fraudulent financial statements. **Expert Systems with Applications**, [S.l.], v. 32, n. 4, p. 995-1003, May 2007.

LEE, T. A.; INGRAM, R. W.; HOWARD, T. P. The difference between earnings and operating cash flow as an indicator of financial reporting fraud. **Contemporary Accounting Research**, Toronto, v. 16, n. 4, p. 749-786, Winter 1999.

LEE, T. H.; GLOECK, J. D.; PALANIAPPAN, A. K. The audit expectation gap: an empirical study in Malaysia. **Southern African Journal of Accountability and Auditing Research**, [S.l.], v. 7, p. 1-15, 2007.

LENNOX, C.; PITTMAN, J. A. Big five audits and accounting fraud. **Contemporary Accounting Research**, Toronto, v. 27, n. 1, p. 209-247, Spring 2010.

LIN, J. W.; HWANG, M. I.; BECKER, J. D. A fuzzy neural network for assessing the risk of fraudulent financial reporting. **Managerial Auditing Journal**, Bradford, v. 18, n. 8, p. 657-665, 2003.

MILLER, G. S. The press as a watchdog for accounting fraud. **Journal of Accounting Research**, Chicago, v. 44, n. 5, p. 1001-1033, Dec. 2006.

OGUT, H. et al. Prediction of financial information manipulation by using support vector machine and probabilistic neural network. **Expert Systems With Applications**, Tarrytown, NY, v. 36, n. 3, p. 5419-5423, Apr. 2009.

PEDNEAULT, S. **Fraud 101**. 3rd ed. Hoboken: John Wiley & Sons, 2009.

PERSONS, O. S. Using financial statement data to identify factors associated with fraudulent financial reporting. **Journal of Applied Business Research**, Littleton, v. 11, n. 3, p. 38-46, June 1995.

POIRIER, D. J. Partial observability in bivariate probit models. **Journal of Econometrics**, Amsterdam, v. 12, n. 2, p. 209-217, Feb. 1980.

POWERS, D. A.; XIE, Y. **Statistical methods for categorical data analysis**. San Diego: Academic Press, 2000.

SKOUSEN, C. J.; WRIGHT, C. J. Contemporaneous risk factors and the prediction of financial statement fraud. **Journal of Forensic Accounting**, Tampa, Florida, v. 9, n. 1, p. 37-62, 2008.

SPATHIS, C. Detecting false financial statements using published data: some evidence from Greece.

Managerial Auditing Journal, Bradford, v. 17, n. 4, p. 179-191, 2002.

SUMMERS, S. L.; SWEENEY, J. T. Fraudulently misstated financial statements and insider trading: an empirical analysis. **The Accounting Review**, Sarasota, v. 73, n. 1, p. 131-146, Jan. 1998.

WANG, T. Y. **Corporate securities fraud: insights from a new empirical framework**. Oxford, 2011. (Journal of Law, Economics, and Organization). No prelo.