

ẢNH HƯỞNG CỦA PHƯƠNG PHÁP CHỌN MẪU ĐỐI VỚI KẾT QUẢ DỰ BÁO KHÓ KHĂN TÀI CHÍNH CHO CÁC CÔNG TY NIÊM YẾT

Vũ Thị Loan

Đại học Kinh tế, Đại học Quốc gia Hà Nội

Email: loanvu.kttn@gmail.com

Đinh Hồng Linh

Đại học Kinh tế và Quản trị kinh doanh – Đại học Thái Nguyên

Email: dhlinh23@gmail.com

Nguyễn Thu Nga

Đại học Kinh tế và Quản trị kinh doanh – Đại học Thái Nguyên

Email: thungadhkt@gmail.com

Ngày nhận: 28/02/2019

Ngày nhận bản sửa: 08/7/2019

Ngày duyệt đăng: 05/10/2019

Tóm tắt:

Nghiên cứu này được tiến hành nhằm xây dựng mô hình dự báo khó khăn tài chính cho các công ty niêm yết trên thị trường chứng khoán Việt Nam có tính đến tác động của phương pháp chọn mẫu đến kết quả dự báo của mô hình. Các công ty gặp khó khăn tài chính là các công ty bị hủy niêm yết bắt buộc do thua lỗ trong ba năm liên tiếp hoặc lỗ lũy kế vượt quá vốn điều lệ. Sử dụng cùng kỹ thuật dự báo Binary Logistic, ba mô hình được xây dựng tương ứng với ba trường hợp số lượng công ty không gặp khó khăn tài chính bằng, gấp đôi và gấp ba lần số lượng công ty khó khăn tài chính trong mẫu nghiên cứu. Kết quả phân tích cho thấy, khi số lượng công ty không gặp khó khăn tài chính tài chính tăng dần thì khả năng dự báo chính xác của mô hình càng cao.

Từ khóa: Dự báo, khó khăn tài chính, mô hình Logistic, phương pháp chọn mẫu.

Mã JEL: G00, G11, G17.

Impacts of sampling technique on the performance of a financial distress prediction model

Abstract:

This study aims to construct a financial distress prediction model for listed firms on Vietnam Stock Exchange that accounts for any effects of sampling technique on the performance of the model. A firm is considered financially distressed when it is required to be delisted by the Vietnam Stock Exchange because of suffering three consecutive years or having its accumulated losses over its charter capital. Using binary logistic as a classifier, there are three models established corresponding with three sample sizes. The analysis shows that the higher the number of non financially distressed firms, the greater the model's prediction accuracy level.

Keywords: Prediction, financial distress, binary logistic, sampling technique.

JEL code: G00, G11, G17.

1. Giới thiệu

Khái niệm khó khăn tài chính (KKTC) được đưa ra lần đầu tiên bởi Beaver (1966) để mô tả tình trạng doanh nghiệp thiếu tiền để thanh toán cho các nghĩa vụ nợ mà hậu quả là phải vay tiền ngân hàng, bán tài sản công ty hay tệ nhất là đứng trên bờ vực phá sản. Vì vậy, Altman (1968) trong nghiên cứu về mô hình chấm điểm Z-score khẳng định rằng phá sản doanh nghiệp chính là một định nghĩa chính thức của khó khăn tài chính. Khó khăn tài chính, trong các nghiên cứu sau này, được nhận biết tách rời với phá sản doanh nghiệp do tình trạng khó khăn tài chính của doanh nghiệp không phải lúc nào cũng dẫn đến tình trạng phá sản (Pindado & cộng sự, 2008). Vì vậy, một mô hình dự báo trở nên rất quan trọng trong việc giúp công ty tránh khỏi tình trạng phá sản (Santoso, 2018). Theo cách tiếp cận này, một công ty gặp khó khăn tài chính nếu: (1) dòng tiền thuần tạo ra nhỏ hơn lãi phải trả và giá trị thị trường của công ty sụt giảm liên tiếp trong hai năm (Asquith & cộng sự, 1994); (2) công ty chịu sự kiểm soát đặc biệt của Sở giao dịch chứng khoán (Xu & cộng sự, 2015; Geng & cộng sự, 2014); (3) cổ phiếu bị hủy hoặc tạm ngừng giao dịch trên thị trường chứng khoán (Lin & cộng sự, 2014); (4) hệ số mô tả khả năng thanh toán lãi vay nhỏ hơn 0,7 (Bhattacharjee & Han, 2014).

Khái niệm dự báo khó khăn tài chính gắn liền với thuật ngữ cảnh báo sớm (early warning) được hiểu là việc nhận biết tình trạng khó khăn tài chính của một chủ thể trong tương lai từ các chỉ báo trong quá khứ và hiện tại. Việc dự báo khó khăn tài chính được bắt nguồn từ lý luận cho rằng hoàn toàn có thể phát hiện các dấu hiệu về khó khăn tài chính của một doanh nghiệp trước khi doanh nghiệp đó chính thức lâm vào tình trạng này. Dự báo khó khăn tài chính đã trở thành vấn đề trọng tâm trong tài chính doanh nghiệp và việc xây dựng các mô hình dự báo khó khăn tài chính đã thu hút rất nhiều sự quan tâm của các nhà nghiên cứu trên toàn thế giới. Các mô hình được xây dựng đã sử dụng một hệ thống các biến dự báo phong phú, các kỹ thuật dự báo đa dạng từ truyền thống đến hiện đại không nằm ngoài mục đích nâng cao độ chính xác trong dự báo. Bên cạnh các biến và kỹ thuật dự báo, do số lượng các công ty gặp khó khăn tài chính chiếm tỷ lệ khá nhỏ trong tổng thể nên việc lựa chọn cỡ mẫu cũng sẽ ảnh hưởng đáng kể đến kết quả dự báo. Có thể thấy rằng, đánh giá về tác động của phương pháp chọn mẫu tới độ chính xác trong dự báo đã được đưa ra từ khá lâu, trong nghiên cứu của Zmijewski (1984) nhưng chủ đề này cũng chưa nhận được sự quan tâm thích đáng trong các nghiên cứu có liên quan. Chính vì

vậy, nghiên cứu này được tiến hành nhằm xây dựng một mô hình dự báo khó khăn tài chính cho các công ty niêm yết trên thị trường chứng khoán Việt Nam có tính đến sự ảnh hưởng của các phương pháp chọn mẫu. Trong đó, một công ty được nhận biết là khó khăn tài chính khi công ty bị hủy niêm yết bắt buộc theo yêu cầu của Ủy ban Chứng khoán Nhà nước do những yếu kém trong hoạt động kinh doanh của mình. Từ đó, kết quả nghiên cứu có thể giúp các nhà quản trị công ty, các nhà đầu tư và các đối tượng có liên quan đưa ra các quyết định phù hợp để nâng cao hiệu quả hoạt động của công ty cũng như hạn chế rủi ro trong kinh doanh.

2. Tổng quan nghiên cứu

2.1. Các yếu tố của một mô hình dự báo khó khăn tài chính

Thuật ngữ “cảnh báo sớm” (Early warning) đã trở nên phổ biến trong kinh tế vĩ mô, quản trị kinh doanh và “cảnh báo sớm” về tình trạng khó khăn tài chính trở thành một đề tài trọng tâm trong quản lý tài chính doanh nghiệp cũng như ngày càng thu hút nhiều nghiên cứu của các học giả trên toàn thế giới. Theo Lin & cộng sự (2014), độ chính xác của một mô hình dự báo phụ thuộc vào việc lựa chọn các biến dự báo và kỹ thuật dự báo.

Các mô hình sử dụng kỹ thuật truyền thống được biết với các đại diện là Beaver (1966), Altman (1968) và sau đó là Ohlson (1980). Năm 1966, Beaver là người tiên phong trong việc đưa ra mô hình dự báo khó khăn tài chính sử dụng các chỉ số tài chính, gọi là mô hình hồi quy đơn biến. Altman (1968) đã giới thiệu kỹ thuật phân tích biệt số nhằm dự báo khó khăn tài chính với mô hình Z-score. Tuy nhiên, sự phổ biến của mô hình biệt số sau đó bị suy giảm do có sự xuất hiện của kỹ thuật phân tích hồi quy Logistic do Ohlson (1980) xây dựng. Mô hình Logistic ước tính xác suất xảy ra khó khăn tài chính của doanh nghiệp mà không kèm theo các yêu cầu về phân phối chuẩn và phương sai bằng nhau đối với dữ liệu như trong mô hình Z-score. Trong các nghiên cứu gần đây, việc dự báo khó khăn tài chính ngày càng được hoàn thiện với việc áp dụng các thuật toán máy học hiện đại như mô hình cây quyết định Decision Tree, mạng thần kinh nhân tạo Neural Network, mô hình máy hỗ trợ vector SVM (Support Vector Machine).

Bên cạnh kỹ thuật dự báo, biến dự báo trong các mô hình khá đa dạng và có thể được chia thành 3 nhóm: chỉ số tài chính, chỉ số kinh tế vĩ mô và các biến thị trường. Nhóm các chỉ số tài chính, tính toán từ hệ thống báo cáo tài chính của doanh nghiệp được

sử dụng trong hầu hết các nghiên cứu vì chúng phản ánh khá đầy đủ các khía cạnh tài chính của công ty: khả năng thanh toán, khả năng sinh lời, cấu trúc vốn,... của doanh nghiệp. Bên cạnh đó, các biến số thị trường cũng trở nên phổ biến trong mô hình dự báo vì chúng chứa đựng các ước tính về thu nhập trong tương lai của doanh nghiệp (Beaver & cộng sự, 2005). Các biến thị trường có thể là giá cổ phiếu, tỷ suất lợi nhuận của chứng khoán và chỉ số vốn hóa thị trường (Beaver & cộng sự, 2005; Tinoco & cộng sự, 2013). Do số lượng các công ty gặp khó khăn tài chính thường cao trong giai đoạn khủng hoảng kinh tế nên các biến số kinh tế vĩ mô cũng có vai trò quan trọng trong dự báo khó khăn tài chính (Koopman & Lucas, 2005). Các chỉ số kinh tế vĩ mô có thể là lãi suất, chu kỳ kinh tế (Bhattacharjee & Han, 2014), GDP, cung tiền và chỉ số giá tiêu dùng CPI (Alifiah, 2014).

Ở Việt Nam, các nghiên cứu có liên quan thường gắn với việc dự báo rủi ro phá sản hay chấm dứt tín dụng trong các ngân hàng thương mại, hầu hết là vận dụng hay điều chỉnh mô hình chấm điểm Z-score của Altman. Có thể kể đến nghiên cứu của Hay Sinh (2013) để ước tính xác suất phá sản của Công ty cổ phần Công nghiệp Cao su Miền Nam dựa trên 4 chỉ số tài chính của công ty này; nghiên cứu của Đào Thị Thanh Bình (2013) cũng sử dụng phương pháp tiếp cận của Altman (1968) để xây dựng mô hình định mức tín dụng cho các công ty sản xuất ở Việt Nam.

2.2. Ảnh hưởng của phương pháp chọn mẫu đối với mô hình dự báo khó khăn tài chính

Trong một mô hình dự báo, bên cạnh việc lựa chọn các kỹ thuật hay thuật toán sử dụng, các nghiên cứu thực nghiệm cũng nhận thấy vai trò của phương pháp chọn mẫu trong việc nâng cao tính chính xác của một mô hình dự báo. Zmijewski (1984) trong nghiên cứu đầu tiên liên quan đến chủ đề này đã phân chia phương pháp chọn mẫu thành hai phần: chọn mẫu phân tầng ngẫu nhiên và phương pháp chọn mẫu hoàn toàn.

Phương pháp chọn mẫu phân tầng ngẫu nhiên được sử dụng khi mẫu nghiên cứu bao gồm tất cả các công ty khó khăn tài chính và chỉ một phần các công ty không gặp khó khăn tài chính. Để lựa chọn vào trong mẫu, các công ty khó khăn tài chính và không khó khăn tài chính thường có sự tương đồng về ngành nghề hoặc quy mô công ty. Chọn mẫu phân tầng ngẫu nhiên được sử dụng rộng rãi trong các nghiên cứu được tiến hành bởi Beaver (1966); Altman (1968); Geng & cộng sự (2014). Như vậy,

tỷ lệ các công ty gặp khó khăn tài chính trong mẫu nghiên cứu thường lớn hơn nhiều so với tỷ lệ của các công ty này trong tổng thể. Vì thế, phương pháp chọn mẫu này có ưu điểm là khắc phục được tình trạng số công ty khó khăn tài chính quá nhỏ so với tổng thể (Shaonan & cộng sự, 2015). Tuy nhiên, tỷ lệ công ty khó khăn tài chính trong mẫu lớn có thể dẫn tới việc các ước lượng của mô hình dự báo không đảm bảo sự chính xác (Zmijewski, 1984; Shaonan & cộng sự, 2015).

Ngược lại với phương pháp chọn mẫu phân tầng ngẫu nhiên, phương pháp chọn mẫu toàn bộ là phương pháp lựa chọn tất cả các công ty không gặp khó khăn tài chính và các công ty khó khăn tài chính vào trong mẫu nghiên cứu. Chẳng hạn, Ohlson (1980) trong nghiên cứu của mình đã lựa chọn tất cả 2.050 công ty không khó khăn tài chính và 105 công ty khó khăn tài chính để xây dựng một mô hình dự báo khó khăn tài chính. Phương pháp chọn mẫu toàn bộ cũng được sử dụng khá rộng rãi trong các nghiên cứu sau này, chẳng hạn nghiên cứu của Bharath & Shumway (2008) hay Kim & Sohn (2010). Phương pháp chọn mẫu này có khả năng nâng cao tính chính xác của các ước lượng do tỷ lệ các công ty khó khăn tài chính trong mẫu tương đồng với tổng thể (Ohlson, 1980; Zmijewski, 1984). Tuy nhiên, phương pháp này đòi hỏi khối lượng tính toán khá lớn và có thể gây ra hiện tượng mất cân bằng dữ liệu (class imbalance) do các lớp dữ liệu (khó khăn tài chính và không khó khăn tài chính) quá chênh lệch và từ đó giảm bớt chất lượng của mô hình dự báo.

Như vậy, có thể thấy rằng, hai phương pháp chọn mẫu cơ bản khác nhau về tỷ lệ các công ty khó khăn tài chính trong mẫu so với tỷ lệ này trong tổng thể. Nếu như phương pháp phân tầng ngẫu nhiên có thể dẫn tới sự kém chính xác trong dự báo thì phương pháp chọn mẫu toàn bộ có khả năng dẫn đến hiện tượng mất cân bằng dữ liệu. Từ đó dẫn đến câu hỏi nghiên cứu về khả năng cải thiện mô hình dự báo khó khăn tài chính nếu như tăng số lượng công ty không gặp khó khăn tài chính khi sử dụng phương pháp chọn mẫu phân tầng ngẫu nhiên.

3. Phương pháp nghiên cứu

3.1. Mục tiêu nghiên cứu

Nghiên cứu được thực hiện với mục tiêu đánh giá sự thay đổi của phương pháp chọn mẫu hay cỡ mẫu được lựa chọn tới kết quả của mô hình dự báo của các công ty niêm yết trên thị trường chứng khoán Việt Nam. Trong đó, công ty gặp khó khăn tài chính là công ty bị hủy niêm yết bắt buộc theo yêu cầu của Ủy ban Chứng khoán Nhà nước. Công ty bị hủy

Bảng 1: Các biến dự báo trong mô hình

Nhóm	Mô tả
Khả năng thanh khoản	X ₁ Tổng nợ/Tổng tài sản
	X ₂ Tài sản ngắn hạn/Nợ ngắn hạn
	X ₃ (Tài sản ngắn hạn – Hàng tồn kho)/Nợ ngắn hạn
	X ₄ Nợ phải trả/Vốn chủ sở hữu
	X ₅ Nợ ngắn hạn/Tổng tài sản
	X ₆ Lưu chuyển thuần từ hoạt động kinh doanh/Nợ ngắn hạn
	X ₇ EBIT/Chi phí lãi vay
Khả năng sinh lời	X ₈ (Doanh thu- Chi phí bán hàng)/Doanh thu
	X ₉ Lợi nhuận ròng/Doanh thu
	X ₁₀ LN trước thuế/Tổng tài sản
	X ₁₁ LN ròng/Tổng tài sản
	X ₁₂ LN ròng/Tài sản ngắn hạn
	X ₁₃ LN ròng/Tài sản cố định
	X ₁₄ LN ròng/Vốn chủ sở hữu
Khả năng hoạt động	X ₁₅ LN từ hoạt động chính/Tổng tài sản
	X ₁₆ Doanh thu bán hàng/Tài sản lưu động
	X ₁₇ Doanh thu bán hàng/Tài sản cố định
	X ₁₈ Chi phí từ hoạt động chính/Hàng tồn kho
	X ₁₉ LN từ hoạt động chính/Phải thu
	X ₂₀ Giá vốn hàng bán/Phải trả
Khả năng tăng trưởng	X ₂₁ Tốc độ tăng trưởng lợi nhuận từ hoạt động chính
	X ₂₂ Tốc độ tăng trưởng tài sản
	X ₂₃ Tốc độ tăng trưởng lợi nhuận ròng
Cấu trúc	X ₂₄ Tài sản ngắn hạn/Tổng tài sản
	X ₂₅ Tài sản cố định/Tổng tài sản
	X ₂₆ Vốn chủ sở hữu/Tài sản dài hạn
	X ₂₇ Nợ ngắn hạn/Tổng nợ phải trả
Khả năng tăng vốn	X ₂₈ LN ròng/Số lượng cổ phiếu lưu hành
	X ₂₉ Tài sản ròng/Số lượng cổ phiếu lưu hành
	X ₃₀ Sự tăng lên của tiền và tương đương tiền/Số lượng CP lưu hành
	X ₃₁ Dự trữ vốn/Số lượng CP lưu hành
Sở hữu nhà nước	X ₃₂ Sở hữu nhà nước

Nguồn: Geng & cộng sự (2014).

niêm yết bắt buộc khi gặp những vấn đề nghiêm trọng về khả năng thanh toán các khoản nợ khi bị lỗ liên tiếp 3 năm hoặc số lỗ lũy kế đã vượt trên mức vốn điều lệ, theo quy định tại Điều 60 Nghị định 58/2012/NĐ-CP của Chính phủ ban hành ngày 20/07/2012.

Trong khoảng thời gian nghiên cứu từ 2009 đến 2017 có 68 công ty bị hủy niêm yết bắt buộc. Nhằm phục vụ cho mục tiêu nghiên cứu, các dữ liệu được thu thập theo quy mô tăng dần tương ứng với 3 mô

hình khác nhau:

- Mô hình 1.1: số lượng các công ty khó khăn tài chính (68 công ty) bằng số các công ty không khó khăn tài chính (68 công ty).

- Mô hình 1.2: số lượng các công ty khó khăn tài chính (68 công ty) bằng 1/2 lần số các công ty không khó khăn tài chính (136 công ty).

- Mô hình 1.3: số lượng các công ty khó khăn tài chính (68 công ty) bằng 1/3 lần số các công ty không khó khăn tài chính (204 công ty).

3.2. Thước đo, biến số

Trong mô hình dự báo, biến phụ thuộc thể hiện tình trạng khó khăn tài chính và nhận hai giá trị như sau: $Y = 1$ nếu công ty gặp khó khăn tài chính; $Y = 0$ nếu công ty không gặp khó khăn tài chính. Các biến độc lập, đóng vai trò dự báo khó khăn tài chính được lựa chọn là 31 chỉ số được sử dụng trong nghiên cứu của Geng & cộng sự (2014) vì các chỉ số này phản ánh đầy đủ các khía cạnh tài chính của doanh nghiệp như khả năng thanh khoản, khả năng sinh lời, khả năng hoạt động, khả năng tăng trưởng và cơ cấu vốn của doanh nghiệp. Tác giả bổ sung biến (X_{32}) nhằm đánh giá vai trò của vốn nhà nước trong cấu trúc vốn của doanh nghiệp tới khả năng gặp khó khăn tài chính của doanh nghiệp. Các biến lựa chọn trong các mô hình được mô tả trong bảng 1.

3.3. Phương pháp phân tích

Dữ liệu thu thập sẽ được xử lý và phân tích sử dụng mô hình Logit nhị thức (Binary Logistic). Mô hình Logit nhị thức là mô hình hồi quy đặc biệt khi biến phụ thuộc là một biến nhị phân chỉ nhận hai giá trị 0 và 1. Mô hình hồi quy dùng để dự đoán xác suất để xảy ra một sự việc dựa vào thông tin từ các biến độc lập trong mô hình. Mô hình này có ưu điểm là không gắn với yêu cầu về phân phối chuẩn và phương sai bằng nhau của biến dự báo vì thế mô hình đã được áp dụng phổ biến trong các nghiên cứu của Ohlson (1980), Hua & cộng sự (2007), Alifiah (2014).

Mô hình Logit có dạng:

$$P = \frac{1}{1 + e^{-(\alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}$$

Trong đó P nhận giá trị 1 nếu công ty khó khăn tài chính và giá trị 0 khi công ty có tình hình tài chính lành mạnh. Xác suất càng lớn hơn 0,5 thì nguy cơ khó khăn tài chính càng cao và ngược lại. Mô hình này được phân tích dưới sự hỗ trợ bởi phần mềm SPSS 23.0.

Khi sử dụng mô hình để dự báo khó khăn tài

chính, các bước phân tích được tiến hành bao gồm: kiểm tra độ phù hợp của mô hình, kiểm định ý nghĩa của các hệ số và kiểm định độ phù hợp tổng quát của mô hình. Độ phù hợp tổng quát của mô hình được dựa trên chỉ tiêu -2LL (-2log likelihood), thước đo này càng nhỏ càng tốt. Mô hình này cũng đòi hỏi kiểm định giả thuyết hệ số hồi quy là khác không. Trong phần mềm SPSS 23.0, mức ý nghĩa quan sát đưa ra trong bảng Omnibus Tests of Model Coefficients để quyết định bác bỏ hay chấp nhận H_0 . Mô hình cũng được coi là phù hợp nếu kiểm định Hosmer & Lemeshow không có ý nghĩa. Mô hình dự báo tốt đến đâu được thể hiện qua bảng phân loại (classification table) trong đó so sánh trị số thực và trị số dự đoán cho từng biểu hiện và tỷ lệ dự đoán đúng sự kiện.

4. Phân tích kết quả nghiên cứu

4.1. Kết quả phân tích sự phù hợp của mô hình

Khi thực hiện các kiểm định cần thiết, mô hình được coi là phù hợp nếu giá trị -2 Log likelihood tính toán được nhỏ, kiểm định Omnibus đảm bảo mức ý nghĩa và kiểm định Hosmer & Lemeshow không có ý nghĩa. Kết quả kiểm định trên bảng 2 cho biết, cả ba mô hình đều có độ tin cậy và phù hợp do các kiểm định cần thiết đều được thỏa mãn.

4.2. Kết quả đánh giá khả năng dự báo của mô hình

Sử dụng kỹ thuật dự báo Binary Logistic để dự báo một công ty vào nhóm khó khăn tài chính ($Y=1$) hoặc không khó khăn tài chính ($Y=0$) dựa vào xác suất tính toán được, độ chính xác trong dự báo trong cả ba mô hình thể hiện trong bảng 3. Bảng này cho biết độ chính xác trong dự báo tăng dần từ mô hình 1.1 cho đến mô hình 1.3. Nói cách khác, khi số lượng công ty không gặp khó khăn tài chính trong mẫu tăng lên thì mô hình dự báo có độ chính xác càng cao. Hay, khi tỷ lệ công ty gặp khó khăn tài chính trong mẫu giảm dần thì mô hình càng dự báo đúng. Độ chính xác của mô hình còn được đánh giá bằng thống kê về sai lầm trong dự báo, trong đó sai

Bảng 2: Kết quả kiểm định mô hình Logit

	Kiểm định Omnibus		-2 Log likelihood	Kiểm định Hosmer and Lemeshow	
	Chi-square	Sig.		Chi-square	Sig.
Mô hình 1.1	101,45	0,00	42,7	1,58	0,99
Mô hình 1.2	93,54	0,00	64,2	8,35	0,40
Mô hình 1.3	87,66	0,00	65,2	2,67	0,95

Nguồn: Kết quả phân tích trên SPSS 23.0.

Bảng 3: Kết quả dự báo khó khăn tài chính

Mô hình	Độ chính xác (%)	Sai lầm loại I (%)	Sai lầm loại II (%)
Mô hình 1.1	78,68	29,41	13,20
Mô hình 1.2	90,70	17,60	5,10
Mô hình 1.3	93,01	11,76	5,40

Nguồn: Kết quả phân tích trên SPSS 23.0.

lầm loại I là sai lầm khi một công ty gặp khó khăn tài chính được dự báo không gặp khó khăn tài chính và sai lầm loại II là sai lầm khi một công ty không khó khăn tài chính được xếp vào nhóm gặp khó khăn tài chính. Kết quả dự báo cho thấy, sai lầm loại I và loại II cũng giảm dần từ mô hình 1.1 cho tới mô hình 1.3. Như vậy, mô hình có khả năng dự báo với độ chính xác cao nhất là mô hình có số lượng công ty không gặp khó khăn tài chính lớn gấp 3 lần công ty gặp khó khăn tài chính.

Từ 32 biến độc lập ban đầu, sử dụng phương pháp LR Backward Stepwise, các biến được loại dần căn cứ vào xác suất của trị thống kê Likelihood-Ratio để tìm ra các biến có khả năng phân biệt tốt nhất. Kết quả cho thấy, khi số lượng công ty không khó khăn tài chính trong mẫu tăng lên, các biến có khả năng phân biệt được giữ lại trong mô hình cũng nhiều hơn. Bảng 4 cho biết, mô hình 1.1 giữ lại 9 biến độc lập trong khi số biến độc lập trong hai mô hình 1.2 và 1.3 lần lượt là 11 và 17. Hệ số tương quan của

các biến độc lập trong từng mô hình cũng được mô tả trên bảng 4. Trong khi đó, tổng hợp các biến được loại ra khỏi 3 mô hình được thể hiện trong phụ lục 01.

Bảng 4 cho biết, mặc dù các biến độc lập là không hoàn toàn giống nhau, nhìn chung cả ba mô hình đều nhấn mạnh tầm quan trọng của khả năng thanh toán (Tổng nợ/Tổng tài sản, Nợ ngắn hạn/Tổng tài sản), khả năng sinh lời (Lợi nhuận ròng/Tổng tài sản, Lợi nhuận ròng/Tài sản ngắn hạn, Lợi nhuận ròng/Tài sản cố định); khả năng tăng vốn (Tài sản ròng/Số lượng cổ phiếu lưu hành). Đặc biệt, biến mô tả sở hữu nhà nước (X_{32}) có ý nghĩa thống kê và khả năng phân biệt trong mô hình 1.3 trong khi biến này đều bị loại bỏ trong mô hình 1.1 và 1.2. Hệ số tương quan của biến X_{32} cho thấy, doanh nghiệp có sở hữu vốn của Nhà nước thì ít có khả năng gặp khó khăn tài chính. Biến X_5 (Nợ ngắn hạn/tổng tài sản) có hệ số tương quan dương lớn nhất trong cả ba mô hình. Ngoài ra, hệ số tương quan của các biến có ý nghĩa

Bảng 4: Hệ số tương quan của các biến trong mô hình

	Mô hình 1.1		Mô hình 1.2		Mô hình 1.3	
	Biến	Hệ số	Biến	Hệ số	Biến	Hệ số
1	X_1	32,39	X_1	21,80	X_1	17,99
2	X_5	35,20	X_4	0,32	X_4	0,23
3	X_6	-1,52	X_5	26,06	X_5	18,42
4	X_9	0,45	X_7	-0,09	X_7	-0,01
5	X_{11}	26,13	X_{12}	-5,32	X_9	1,29
6	X_{12}	-12,76	X_{13}	-2,37	X_{12}	-7,27
7	X_{13}	-6,46	X_{14}	0,77	X_{13}	-2,57
8	X_{15}	-6,49	X_{19}	-0,36	X_{14}	1,21
9	X_{27}	23,14	X_{20}	-0,55	X_{17}	-0,13
10			X_{27}	16,80	X_{19}	-0,71
11			X_{29}	0,04	X_{21}	0,05
12					X_{22}	0,32
13					X_{23}	0,04
14					X_{27}	11,03
15					X_{29}	0,05
16					X_{30}	-0,14
17					X_{32}	-1,31

Nguồn: Kết quả phân tích trên SPSS 23.0.

trong cả 3 mô hình đều mang dấu phù hợp với suy đoán ban đầu về mối quan hệ giữa biến độc lập và biến phụ thuộc.

5. Kết luận

Nghiên cứu này được tiến hành nhằm xây dựng một mô hình dự báo khó khăn tài chính cho các công ty niêm yết trên thị trường chứng khoán Việt Nam có tính đến sự tác động của các phương pháp chọn mẫu khác nhau. Từ 32 biến dự báo ban đầu, các biến được lựa chọn đã giảm xuống đáng kể nhằm bảo đảm độ tin cậy và mức ý nghĩa của các biến này. Kết quả nghiên cứu cho thấy các biến liên quan đến hệ số nợ trên tổng tài sản, khả năng sinh lời, khả năng hoạt động và khả năng tăng vốn đóng vai trò quan trọng trong việc dự báo khó khăn tài chính của các công ty niêm yết. Ngoài ra, biến độc lập mô tả sự có mặt của vốn Nhà nước trong cơ cấu vốn của doanh nghiệp cũng có khả năng dự báo khó khăn tài chính trong mô hình có số công ty không gặp khó khăn tài chính nhiều nhất (mô hình 1.3).

Kết quả phân tích các mô hình khác nhau cho thấy, việc sử dụng các cỡ mẫu khác nhau về số lượng các công ty không khó khăn tài chính có ảnh hưởng đến kết quả dự báo của mô hình. Dù số lượng công ty khó khăn tài chính là như nhau nhưng mô hình càng có nhiều công ty không khó khăn tài chính thì độ chính xác trong dự báo càng cao. Nói cách khác, tỷ lệ các công ty khó khăn tài chính trong mẫu nghiên cứu gần với tỷ lệ đó của tổng thể thì khả năng dự báo của mô hình càng tốt. Kết hợp với nhận xét về số lượng biến độc lập trong mô hình có thể thấy mô hình bao gồm càng nhiều công ty không khó khăn tài chính trong mẫu nghiên cứu càng đòi hỏi nhiều biến dự báo nhưng lại có kết quả dự báo càng cao và ngược lại. Tuy nhiên, kết quả nghiên cứu cũng đặt ra hướng nghiên cứu mới về việc độ chính xác của mô hình có tiếp tục tăng lên nếu tỷ lệ các công ty không khó khăn tài chính trong mẫu nghiên cứu rất gần với tổng thể như trong phương pháp lấy mẫu toàn bộ hay không.

Phụ lục 01: Tổng hợp các biến dự báo được loại khỏi các mô hình

TT	Mô hình 1.1		Mô hình 1.2		Mô hình 1.3	
	Biến	Mức ý nghĩa	Biến	Mức ý nghĩa	Biến	Mức ý nghĩa
1	X ₂	0,78	X ₂	0,26	X ₂	0,29
2	X ₃	0,72	X ₃	0,35	X ₃	0,14
3	X ₄	0,45	X ₆	0,49	X ₆	0,20
4	X ₇	0,73	X ₆	0,76	X ₈	0,80
5	X ₈	0,61	X ₉	0,35	X ₁₀	0,81
6	X ₁₀	0,91	X ₁₀	0,96	X ₁₁	0,70
7	X ₁₄	0,06	X ₁₁	0,48	X ₁₅	0,25
8	X ₁₆	0,78	X ₁₅	0,23	X ₁₆	0,80
9	X ₁₇	0,49	X ₁₆	0,13	X ₁₈	0,70
10	X ₁₈	0,65	X ₁₇	0,46	X ₂₀	0,16
11	X ₁₉	0,73	X ₁₈	0,75	X ₂₄	0,22
12	X ₂₀	0,88	X ₂₁	0,93	X ₂₅	1,00
13	X ₂₁	0,05	X ₂₂	0,27	X ₂₆	0,82
14	X ₂₂	0,96	X ₂₃	0,27	X ₂₈	0,88
15	X ₂₃	0,94	X ₂₄	0,16	X ₃₁	0,67
16	X ₂₄	0,77	X ₂₅	0,54		
17	X ₂₅	0,27	X ₂₆	0,76		
18	X ₂₆	0,05	X ₂₈	0,45		
19	X ₂₈	0,50	X ₃₀	0,93		
20	X ₂₉	0,37	X ₃₁	0,72		
21	X ₃₀	0,48	X ₃₂	0,25		
22	X ₃₁	0,90				
23	X ₃₂	0,53				

Nguồn: Kết quả phân tích trên SPSS 23.0.

Tài liệu tham khảo:

- Alifiah, M. (2014), 'Prediction of financial distress companies in the trading and services sector in Malaysia using macroeconomic variables', *Procedia - Social and Behavioral Sciences*, 129, 90 – 98.
- Altman, E. I. (1968), 'Financial Ratios, discriminant analysis and the prediction of corporate bankruptcy', *Journal of finance*, 23, 589-609.
- Asquith, P., Gertner, R. & Scharfstein, D. (1994), 'Anatomy of financial distress: an examination of junk-bond issuers,' *Q J Econ*, 109, 625–58
- Beaver W. (1966), 'Financial ratios as predictors of failures', *Journal of Accounting research*, 4, 71-111.
- Beaver, W. H., McNichols, M. F. & Rhie, J. W. (2005), 'Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy', *Review of Accounting Studies*, 10, 93–122.
- Bharath, S. & Shumay, T. (2008), 'Forecasting Default with the Merton Distance to Default Model', *Rev. Financ. Study*, 21, 1339–1369.
- Bhattacharjee A. & Han J. (2014), 'Financial distress of Chinese firms: Microeconomic, macroeconomic and institutional influences', *China Economic Review*, 30, 244–262.
- Đào Thị Thanh Bình (2013), 'Mô hình xếp hạng tín dụng cho các công ty sản xuất ở Việt Nam', *Tạp chí Kinh tế và Phát triển*, 188, 39-49.
- Geng R., Bose I. & Chen, X. (2014), 'Prediction of financial distress: An empirical study of listed Chinese companies using data mining', *European of Operational Research*, 241, 236-247.
- Hay Sinh (2013), 'Ước tính xác suất phá sản trong thẩm định giá trị doanh nghiệp', *Phát triển & Hội nhập*, 8(18), 52-57.
- Hua, Z., Wang, Y., Xu, X., Zhang, B. & Liang, L. (2007), 'Predicting corporate financial distress based on integration of support vector machine and logistic regression', *Expert Systems with Applications*, 33(2), 434–440.
- Kim, H. S. & Sohn, S.Y. (2010), 'Support vector machines for default prediction of SMEs based on technology credit', *European Journal of Operational Research*, 201, 838–846.
- Koopman, S.J. & Lucas, A. (2005), 'Business and default cycles for credit risk', *Journal of Applied Econometrics*, 20, 311-323
- Lin, F., Liang, D., Yeh, C. & Huang J. (2014), 'Novel feature selection methods to financial distress prediction', *Expert Systems with Applications*, 41(5), 2472–2483.
- Ohlson, D. (1980), 'Financial ratios and the probabilistic prediction of bankruptcy', *Journal of Accounting Research*, 18(1), 109–131
- Pindado, J., Rodrigues, L. & Torre, C. (2008), 'Estimating financial distress likelihood', *Journal of Business Research*, 61, 995–1003
- Santoso, N., Wibowo, W. (2018), 'Financial Distress Prediction using Linear Discriminant Analysis and Support Vector Machine', *J. Phys.: Conf. Ser* 979, 1-7, DOI:10.1088/1742-6596/979/1/012089
- Shaonan, T., Yan, Y. & Guo, H. (2015), 'Variable selection and corporate bankruptcy forecasts', *Journal of Banking & Finance*, 52(C), 89-100
- Tinoco, M. H., Holmes, P. & Wilson, N. (2013), 'Polytomous response financial distress models: The role of accounting, market and macroeconomic variables', *International Review of Financial Analysis*, 1-39, doi:10.1016/j.irfa.2018.03.017
- Xu, K., Zhao, Q. & Bao, X. (2015), 'Study on early warning of enterprise financial distress – based on partial least squares logistic regression', *Acta Oeconomica*, 65, 3–16, DOI: <http://dx.doi.org/10.1556/032.65.2015.S2.2>
- Zmijewski, M. (1984), 'Methodological issues related to the estimation of financial distress prediction models', *Journal of Accounting Research*, 22, 59-82.