

Sử dụng hàm hồi quy Binary Logistic xây dựng mô hình dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container tại Việt Nam

TRƯƠNG QUÝ HÀO*

Tóm tắt

Thông qua thông số đầu tư tàu container tại Việt Nam giai đoạn 1997-2019, bài viết sử dụng hàm hồi quy Binary Logistic nhằm xây dựng mô hình dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam. Kết quả cho thấy, có 5 nhân tố tác động được đưa vào phân tích nhằm dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam, gồm: Suất đầu tư; Suất vay vốn; Khả năng tài chính/hệ số nợ; Chu phí vốn bình quân; và Thời gian vay vốn.

Từ khóa: container, Binary Logistic, rủi ro tín dụng

Summary

Through the data of investment in container in Vietnam during the period 1997-2019, the article employs the Binary Logistic regression to build a credit risk prediction model for the investment in container fleet in Vietnam. The result introduces five influential factors to the aforementioned model, including Invested capital; Borrowed capital; Financial capability/debt ratio; Weighted average cost of capital; Loan period

Keywords: container, Binary Logistic, credit risk

GIỚI THIỆU

Nghi quyết số 36/NQ-TW, ngày 22/10/2018 của Ban chấp hành trung ương Đảng cộng sản Việt Nam khóa XII về Chiến lược phát triển bền vững kinh tế biển Việt Nam đến năm 2030 tầm nhìn đến năm 2045 nêu rõ "đến năm 2030 các ngành kinh tế thuần biển đóng góp khoảng 10% GDP cả nước; kinh tế của 28 tỉnh, thành phố ven biển ước đạt 65%-70% GDP cả nước. Các ngành kinh tế biển phát triển bền vững theo các chuẩn mực quốc tế; kiểm soát khai thác tài nguyên biển trong khả năng phục hồi của hệ sinh thái biển. Đến năm 2045, Việt Nam trở thành quốc gia biển mạnh, phát triển bền vững, thịnh vượng, an ninh, an toàn; kinh tế biển đóng góp quan trọng vào nền kinh tế đất nước, góp phần xây dựng nước ta thành nước công nghiệp hiện đại theo

định hướng xã hội chủ nghĩa; tham gia chủ động và có trách nhiệm vào giải quyết các vấn đề quốc tế và khu vực về biển và đại dương".

Để thực hiện được các mục tiêu này, cùng với các chính sách của Nhà nước, các doanh nghiệp trong lĩnh vực kinh doanh đội tàu cần phải triển khai các hoạt động tàu container cần sự hỗ trợ rất lớn từ phía các tổ chức tín dụng. Tuy nhiên, hoạt động kinh doanh luôn tiềm ẩn nhiều rủi ro. Để hoạt động đầu tư phát triển đội tàu container hiệu quả, cần thiết phải tìm hiểu và phân tích các yếu tố rủi ro nhất là những rủi ro về mặt tín dụng (*Bài viết sử dụng cách viết số thập phân theo chuẩn quốc tế*).

PHƯƠNG PHÁP NGHIÊN CỨU

Hồi quy Binary logistic sử dụng biến phụ thuộc dạng nhị phân để ước lượng xác suất một sự kiện sẽ xảy ra với những thông tin của biến độc lập, mà ta có được (Hoàng Trọng và Chu Nguyễn Mộng Ngọc, 2008). Hai biểu hiện này sẽ được mã hóa thành hai giá trị 0 và 1 và ở dưới dạng này gọi là biến nhị phân.

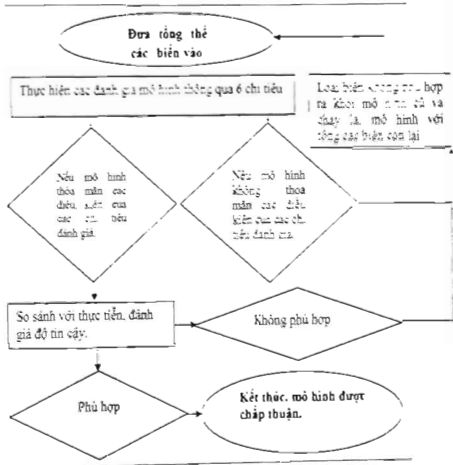
*ThS., Trường Đại học Hàng hải Việt Nam

Ngày nhận bài: 10/12/2019; Ngày phân biên: 05/01/2020; Ngày duyệt đăng: 13/01/2020

BẢNG 1. BẢNG MÃ HOA BIẾN CÁC NHÂN TỐ ẢNH HƯỞNG TỚI RỦI RO TÍN DỤNG TRONG ĐẦU TƯ PHÁT TRIỂN ĐỘI TÀU CONTAINER CỦA VIỆT NAM

STT	Nguyên Nhân	Mã hóa biến
1	Tổng mức đầu tư (Triệu đồng)	X ₁
2	Số lượng đầu tư (Triệu đồng/TEU)	X ₂
3	Số lượng vay vốn (Triệu đồng/TEU)	X ₃
4	Khả năng tài chính/hệ số nợ (%)	X ₄
5	Kiểm nghiệm quản lý khai thác (năm) của chủ tàu	X ₅
6	Uy tín của chủ đầu tư/Mức tín nhiệm (Điểm)	X ₆
7	Chi phí vốn bình quân (Triệu đồng/năm/TEU)	X ₇
8	Chi phí khấu hao (Triệu đồng/năm)	X ₈
9	Thời gian vay vốn (năm)	X ₉

HÌNH: LƯỚI ĐÓNG QUY TRÌNH XÂY DỰNG VÀ LỰA CHỌN MÔ HÌNH DỰ BÁO TỐI ƯU



Để nhận thấy hai biến phụ thuộc của hàm Binary Logistic cho hai kết quả nếu ứng dụng vào mục tiêu dự báo rủi ro là:

- + Kết quả là 0 - tương ứng với trạng thái dự báo "Không có rủi ro tín dụng"
- + Kết quả là 1 - tương ứng với trạng thái dự báo "Có rủi ro tín dụng"

Hàm cho kết quả của biến phụ thuộc như trên hoàn toàn phù hợp với mục tiêu nghiên cứu. Mô hình hàm hồi quy Binary Logistic trong trường hợp đơn giản nhất đó là khi hàm chỉ có một biến độc lập X. Ta có hàm Binary Logistic dạng đơn giản, như sau:

$$P_i = E\left(\frac{Y}{X}\right) = 1 = \frac{e^{B_0 - B_1 X}}{1 + e^{B_0 - B_1 X}} \quad (1)$$

Trong đó: $P_i = E\left(\frac{Y}{X}\right) = P_i (Y=1)$ là xác suất để Y = 1 (xác suất để sự kiện có rủi ro tín dụng xảy ra) khi biến độc lập X có giá trị cụ thể là $X_i[1]$. P_i có giá trị trong khoảng [0;1], với giá trị $P_i \geq 0.5$ sẽ cho kết quả Y = 1; ngược lại với $P_i < 0.5$ sẽ cho kết quả Y = 0.

Căn cứ vào các nhân tố ảnh hưởng tới rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam, tác giả mã hóa biến các nhân tố chính dẫn tới rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam như Bảng 1.

Với 9 biến trên, áp dụng vào hàm Binary Logistic ta có mô hình dự báo với 9 biến như sau:

$$P_i = E(Y) = \frac{\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 - \beta_5 X_5 - \beta_6 X_6 - \beta_7 X_7 - \beta_8 X_8 - \beta_9 X_9}{1 + e^{\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 - \beta_5 X_5 - \beta_6 X_6 - \beta_7 X_7 - \beta_8 X_8 - \beta_9 X_9}} \quad (2)$$

Như vậy, mô hình hàm dự báo được dự kiến gồm 9 biến, với hệ số β tương ứng cho từng biến $\beta_0 - \beta_9$, đã được thiết lập theo công thức (2) ở trên.

Quá trình xây dựng mô hình dự báo tối ưu được kiểm tra đánh giá thông qua các chỉ tiêu đo lường độ tin cậy và mối quan hệ của các biến trong mô hình nhờ phần mềm SPSS với các chỉ tiêu đánh giá cụ thể, gồm:

- Độ phù hợp của mô hình thông qua chỉ tiêu: $-2LL - 2 \log \text{likelihood}$, càng nhỏ càng tốt.

- Kiểm định ý nghĩa các hệ số với chỉ tiêu Wald Chi Square = $[\beta / \text{s.e.}(\beta)]^2 = (B \cdot e) / B$

- Kiểm định độ phù hợp tổng quát, theo chỉ tiêu định Chi - bình phương (X^2) :

$$X^2 = \frac{(O - E)^2}{E}$$

- Kiểm định độ tin cậy của thang đo theo chỉ tiêu Cronbach's Alpha, trong đó độ tin cậy tổng thể ≥ 0.6 , độ tin cậy của từng nhân tố ≥ 0.3

- Kiểm định nhân tố khám phá (EFA) thông qua chỉ tiêu KMO (Kaiser-Meyer-Olkin) với yêu cầu KMO ≥ 0.5

- Phân tích tương quan với hệ số Sig thể hiện tính phụ thuộc của hệ số tương quan giữa các biến theo phép kiểm định với một độ tin cậy cho trước, thường là 95% nên Sig ≤ 0.05

Quá trình kiểm định, lựa chọn biến không phù hợp để đưa ra mô hình tối ưu được thực hiện như Hình.

Nghiên cứu sử dụng bộ dữ liệu về đội tàu container của Cục Hàng hải Việt Nam và Trung tâm Thông tin Tin dụng Quốc gia trong giai đoạn 1997-2019.

Với các giả thuyết các nhân tố khác không ảnh hưởng, số liệu về các nguyên nhân chính ảnh hưởng tới rủi ro tín dụng tại 50 dự án đầu tư tàu container như Bảng 2.

BẢNG 2: BẢNG TỔNG HỢP CÁC THÔNG SỐ ĐẦU TƯ TÀU CONTAINER TẠI VIỆT NAM GIAI ĐOẠN 1997-2019

Tên đầu tư	X1	X2	X3	X4	X5	X6	X7	X8	X9	Y
T1	44.000	109	76	60	7	76	12,0	5,500	5	0
T2	370.000	617	555	80	24	70	21,0	24,667	15	1
T3	330.000	541	487	82	24	70	21,0	22,000	15	1
T4	410.000	404	363	83	24	70	21,0	27,333	10	1
T5	410.000	404	363	78	24	70	19,0	27,333	10	1
T6	320.000	525	472	82	24	70	21,0	21,333	10	1
T7	250.000	595	536	77	24	70	18,0	16,667	12	1
T8	250.000	619	557	75	24	70	15,0	16,667	12	1
T9	300.000	714	643	85	19	70	19,0	30,000	10	1
T10	300.000	714	643	90	19	70	19,0	30,000	10	1
T11	135.000	241	150	67	49	72	9,0	8,000	5	1
T12	125.000	223	150	64	49	72	9,0	8,000	5	1
T13	93.500	119	71	34	10	86	4,0	9,350	5	0
T14	84.200	75	26	40	10	85	4,0	8,420	5	0
T15	217.000	205	82	18	10	86	5,0	14,467	5	0
T16	112.000	93	47	19	10	86	5,0	7,467	7	0
T17	125.000	69	35	20	10	85	5,0	8,333	5	0
T18	125.000	121	36	43	10	87	3,8	12,500	5	0
T19	80.000	113	57	50	3	75	7,0	8,000	5	0
T20	400.000	571	571	90	12	60	19,0	26,667	8	1
T21	170.000	449	449	70	25	65	19,0	11,333	15	1
T22	230.000	885	885	80	25	68	17,0	15,333	13	1
T23	230.000	885	885	85	25	65	15,0	15,333	13	1
T24	80.000	114	57	30	29	85	11,0	8,000	8	0
T25	60.000	104	42	33	29	85	11,0	8,571	6	0
T26	96.000	143	100	65	12	75	12,0	9,600	8	0
T27	120.000	171	120	70	12	75	12,0	12,000	5	0
T28	90.000	160	112	73	12	74	12,0	11,250	6	0
T29	210.000	500	200	50	6	80	12,0	21,000	5	0
T30	250.000	357	179	80	6	80	12,0	25,000	5	0
T31	200.000	328	197	80	6	80	12,0	20,000	5	0
T32	50.000	71	43	85	14	85	18,0	7,143	5	0
T33	75.000	114	57	85	14	85	12,0	7,500	8	0
T34	75.000	115	69	15	15	79	12,0	7,500	7	0
T35	65.600	131	66	10	15	79	11,0	6,560	7	0
T36	70.000	100	50	40	19	70	12,0	7,000	5	0
T37	60.000	200	100	83	19	70	18,0	4,000	10	0
T38	41.000	163	81	80	19	70	10,5	4,100	8	0
T39	220.000	393	354	90	24	65	15,0	14,667	12	1
T40	530.000	491	442	90	24	65	16,0	35,333	15	1
T41	920.000	511	460	90	24	65	15,0	61,333	15	1
T42	154.000	367	257	70	12	67	20,0	10,267	8	0
T43	80.000	141	99	70	12	67	14,0	8,000	6	0
T44	225.900	389	351	78	16	68	6,5	15,060	12	0
T45	238.800	412	329	80	13	68	6,0	15,920	12	1
T46	350.000	603	543	80	12	60	19,0	23,333	12	1
T47	57.000	127	108	60	21	80	5,9	7,125	6	0
T48	86.000	91	59	45	21	80	6,0	8,600	3	0
T49	101.000	106	80	56	21	80	7,8	6,733	4	0
T50	70.000	173	104	70	11	76	10,0	7,000	7	0

KẾT QUẢ NGHIÊN CỨU

Với kết quả thu thập được, tác giả tiến hành nhập liệu vào phần mềm SPSS và sử dụng mô hình hồi quy Binary Logistic như đã phân tích ở trên. Kết quả sau khi thực hiện phân tích đối với 9 biến phụ thuộc như mô hình tại công thức (2) cho thấy, loại 4 biến X1, X5, X6, X8 do không đảm bảo độ tin cậy. Lúc này mô hình hàm dự báo tối ưu mà tác giả đề xuất gồm 5 biến như sau:

$$F_i = E(Y = 1|X_i) = \frac{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5}}{1 + \beta_6 + \beta_7 X_{i6} + \beta_8 X_{i7} + \beta_9 X_{i8} + \beta_{10} X_{i9} + \beta_{11} X_{i10}} \quad (3)$$

Thực hiện kiểm định mô hình với số liệu thu thập được bằng t có kết quả:

Bảng 3 cho thấy, mức ý nghĩa quan sát Sig. = 0.000, nên tác giả có cơ sở bác bỏ giả thuyết B1 = B2 = ... = B7 = 0, kết luận các hệ số của mô hình có ý nghĩa.

Bảng 4 cho thấy, giá trị -2 Log likelihood nhỏ (-2 Log likelihood = 15.4) như vậy chứng tỏ độ phù hợp tốt của mô hình tổng thể.

Bảng 5, 6 cho thấy, trong 30 trường hợp không có xảy ra rủi ro tín dụng, mô hình dự báo đúng 29 trường hợp đạt 96.7% trường hợp; trong số 20 trường hợp có xảy ra rủi ro tín dụng mô hình dự báo đúng 18 trường hợp như vậy. Vậy, tỷ lệ dự báo đúng đạt 90%. Độ chính xác tổng thể là 94%, trong khoa học dự báo độ tin cậy của mô hình như trên là tốt. Tuy nhiên, để tiếp tục đánh giá sự phù hợp của mô hình này, tác giả tiến hành các phân tích EFA.

Phân tích EFA

Kết quả phân tích EFA Bảng 7 cho thấy, hệ số KMO = 0.785 (điều kiện 0.5 < KMO < 1), từ đó cho thấy 5 nhân tố trong mô hình 2 có mức độ phù hợp cao. Hệ số Sig. của kiểm định Bartlett với Sig. = 0.000 (< 0.05), thì các biến quan sát có tương quan với nhau trong tổng thể.

Phân tích tương quan

Kết quả chạy phân tích tương quan (Bảng 8) giữa các biến thông qua phân

Để đảm bảo tính bảo mật thông tin cho các doanh nghiệp đang hoạt động tác giả đổi tên các tàu container và ký hiệu là T.

Nguồn: Cục Hàng hải Việt Nam; Trung tâm Thông tin Tin dụng Quốc gia và sinh viên của tác giả

BẢNG 3: OMNIBUS TESTS OF MODEL COEFFICIENTS

Bước	Chi-square	df	Sig.
Step 1	Step	51.858	.5
	Block	51.858	.5
	Model	51.858	.5

BẢNG 4: MODEL SUMMARY

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	15.444		.646

Nguồn: Tính toán của tác giả

BẢNG 5: CLASSIFICATION TABLE

	Observed	Predicted			
		Y		Percentage Correct	
	0	1			
Step 1	Y	0	29	1	96.7
		1	2	18	90.0
Overall Percentage					94.0

BẢNG 6: VARIABLES IN THE EQUATION

Biến độc lập	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 1	X2	-.022	.020	1.217	1	.270	.978
	X3	.067	.041	2.650	1	.104	1.069
	X4	.015	.085	.031	1	.861	1.015
	X7	-.098	.175	.309	1	.578	.907
	X9	-1.025	.824	1.551	1	.213	.359
	Constant	-.937	5.347	.031	1	.861	.392

BẢNG 7: KMO AND BARTLETT'S TEST

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.795
Bartlett's Test of Sphericity	Approx. Chi-Square	252.738
	df	10
	Sig.	.000

BẢNG 8: CORRELATIONS

Biến		X2	X3	X4	X7	X9	Y
X2	Pearson Correlation	1	.979**	.612**	.650**	.744**	.799**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	50	50	50	50	50	50
X3	Pearson Correlation	.979**	1	.616**	.655**	.780**	.829**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	50	50	50	50	50	50
X4	Pearson Correlation	.612**	.616**	1	.644**	.559**	.583**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	50	50	50	50	50	50
X7	Pearson Correlation	.650**	.655**	.644**	1	.619**	.606**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	50	50	50	50	50	50
X9	Pearson Correlation	.744**	.780**	.559**	.619**	1	.738**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	50	50	50	50	50	50
Y	Pearson Correlation	.799**	.829**	.583**	.606**	.738**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	50	50	50	50	50	50

TÀI LIỆU THAM KHẢO

- Ban Chấp hành Trung ương Đảng (2018). *Nghị quyết số 36/NQ-TW, ngày 22/10/2018 về Chiến lược Phát triển bền vững kinh tế biển Việt Nam đến năm 2030 tầm nhìn đến năm 2045*
- Nguyễn Đình Thọ (2011). *Phương pháp nghiên cứu khoa học trong kinh doanh*. NXB Tài Chính
- Hoàng Trọng - Chu Nguyễn Mộng Ngọc (2008). *Phân tích dữ liệu nghiên cứu với SPSS*. tập 2, NXB Đại học kinh tế TP. Hồ Chí Minh
- Trương Quý Hào (2019). Các nhân tố ảnh hưởng tới rủi ro tín dụng trong đầu tư phát triển tàu container của Việt Nam, *Tạp chí Giao thông vận tải*, số 12/2019

mềm SPSS cho thấy, Sig. giữa biến "Y" và 5 nhân tố đều nhỏ hơn 0.05 chứng tỏ mức độ tương quan giữa 5 biến độc lập và biến phụ thuộc có ý nghĩa thống kê. do đó không loại biến nào ra khỏi mô hình. Như vậy, mô hình tối ưu đủ (gồm 5 nhân tố) chính là mô hình tối ưu đủ điều kiện để tác giả sử dụng làm công cụ dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container tại Việt Nam với độ tin cậy tốt và độ chính xác đạt 94%.

Như vậy, sau quá trình thực nghiệm lựa chọn mô hình tối ưu, kết quả nghiên cứu đưa ra mô hình dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam sử dụng hàm hồi quy Binary Logistic, với 5 biến theo hướng trình dưới đây:

$$\begin{aligned}
 P_i = F(Y = 1|X_j) = \frac{e^{-0.937 - 0.022X_2 + 0.067X_3}}{1 + e^{-0.937 - 0.022X_2 + 0.067X_3}} \\
 + 0.015X_4 - 0.098X_7 - 1.025X_9 \\
 + 0.015X_4 - 0.098X_7 - 1.025X_9 \quad (4)
 \end{aligned}$$

KẾT LUẬN

Việc ứng dụng hàm hồi quy Binary Logistic rất phù hợp cho việc xây dựng mô hình dự báo rủi ro tín dụng trong đầu tư phát triển đội tàu container của Việt Nam. Thực tế tác giả đã xây dựng thành công mô hình dự báo rủi ro tín dụng như công thức (4), từ đó có thể lượng hóa rủi ro trong đầu tư phát triển đội tàu container của Việt Nam, tác giả mong muốn thông qua kết quả, mà mô hình dự báo mang lại giúp các tổ chức tín dụng khi thẩm định dự án đầu tư tàu container sẽ có thêm một công cụ dự báo khả năng xảy ra rủi ro tín dụng đối với phương án tài trợ. Từ đó, có được quyết định đúng đắn giảm thiểu và kiểm soát được rủi ro tín dụng giúp lĩnh vực đầu tư phát triển đội tàu container của Việt Nam phát triển mạnh mẽ và bền vững □