ARTICLE IN PRESS

Advances in Accounting xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Advances in Accounting

journal homepage: www.elsevier.com/locate/adiac

Bankruptcy prediction models' generalizability: Evidence from emerging market economies

Ibrahim Onur Oz[a,*], Can Simga-Mugan[b]

[a] Accounting and Taxation Department, Barney School of Business, University of Hartford, 200 Bloomfield Avenue, West Hartford, CT 06117, USA
[b] School of Business, Izmir University of Economics, Sakarya Caddesi, No: 156, 35330 Balcova, Turkey

1. Introduction

Even as modern researchers and practitioners recognize the critical need for more accurate bankruptcy and distress prediction models, a lack of consensus remains regarding how various proposed models perform in different economic circumstances. In particular, available bankruptcy prediction models might not generalize across economic environments, such as those that mark different nations. By scrutinizing the prediction capability of models across countries, the current study seeks to extend prior literature that tends to investigate prediction models only in relation to developed economies (e.g., Agarwal & Taffler, 2007, 2008; Boritz, Kennedy, & Sun, 2007). But such studies necessarily reflect the unique traits of their samples, suggesting the powerful demand for cross-country analyses of extant models (Altman, Iwanicz-Drozdowska, Laitinen, & Suvas, 2017), across economies that represent diverse settings. Furthermore, some prediction models fail to establish a firm theoretical basis for their financial ratio selections (Charitou, Neophytou, & Charalambous, 2004; Gentry, Newbold, & Whitford, 1985a; Grice & Dugan, 2003; Oz & Yelkenci, 2017), which could imply even greater sample dependence.

To explore existing bankruptcy prediction models' generalizability, and in particular their applicability to emerging economies, this study focuses on five prominent models proposed by Altman (1968), Ohlson (1980), Taffler (1983), Zmijewski (1984), and Shumway (2001). All five of these models originally were derived with samples that came from developed economies, whereas their applicability to emerging economy samples has not been tested. Furthermore, the models originally applied to industrial firms, and the health of such firms is central to the efforts of emerging markets to participate in the global economy (Khanna & Palepu, 2006; Oz & Yelkenci, 2017). In this sense, confirming the generalizability of these models would provide pertinent insights for research but also hold promise for informing practitioners about which prediction models they should adopt.

Some previous research already has established that these prediction models are generalizable in terms of their classification accuracy across different samples (e.g., Grice & Dugan, 2001, 2003; Grice & Ingram, 2001). That is, these studies show that the prediction models can detect company distress accurately, independent of the observation samples. But in addition to testing the generalizability of these prediction models across different samples, it also is necessary to test for re-estimations of the model coefficients (Grice & Dugan, 2003) and confirm the statistical significance of the prediction results (Grice & Dugan, 2001). Few tests of proposed prediction models include these research considerations though. Instead, most research tends to implement specific prediction models for individual country samples, to measure and compare their prediction performance (Kordlar & Nikbakht, 2011; Lifschutz & Jacobi, 2010; Oude, 2013; Pongsatat, Ramage, & Lawrence, 2004), or else apply original versions of the models without examining the statistical validity of their results (Almamy, Aston, & Ngwa, 2016; Chouhan, Chandra, & Gosvami, 2014; Hussain, Ali, Ullah, & Ali, 2014; Malik, Aftab, & Noreen, 2013; Mizan & Hossain, 2014). To extend beyond such considerations, the present study checks the effectiveness of the five models across an economically diverse, multicountry sample.

In addition, this study acknowledges potential differences in the pre- and post-2008 financial crisis periods. By leveraging the impacts of this global economic event, it is possible to examine each model's predictive ability in the aftermath of widespread damage to firm performance (Mihajlov, 2014; Notta & Vlachei, 2014; Oz & Balsari, 2017; Yap, Mohamed, & Chong, 2014). Across these varying economic circumstances, tests of these prediction models can contribute to the generalizability analysis, by demonstrating their capacity—or inability—to detect distress with respect to the financial crisis.

For this study, financial distress is defined as two consecutive years of negative income by existing firms (see also Altman, 1968; DeAngelo & DeAngelo, 1990; Hill, Perry, & Andes, 1996; Li & Sun, 2008; Oz & Yelkenci, 2015). By adopting this measure of financial distress, which represents a uniform outcome variable, this study seeks to validate the classification results more effectively (Oz & Yelkenci, 2017), whereas previous generalizability studies have not included such a measure (Grice & Dugan, 2001, 2003; Grice & Ingram, 2001; Lawrence, Pongstat, & Lawrence, 2015). With a constant firm sample over time, this research also attempts to overcome the bankruptcy proportion problem, as noted by Zmijewski (1984), which arises from the

* Corresponding author.
E-mail addresses: oz@hartford.edu (I.O. Oz), can.mugan@ieu.edu.tr (C. Simga-Mugan).

https://doi.org/10.1016/j.adiac.2018.02.002
Received 28 September 2017; Received in revised form 1 February 2018; Accepted 2 February 2018
0882-6110/

mismatch between the number of bankruptcies in the estimation sample versus the actual (smaller) number of firms that experience bankruptcy in reality.

With this unique approach, the current study reveals that bankruptcy prediction models can produce generalizable predictions for multicountry samples. Despite being proposed in developed economy contexts, they are capable of detecting distressed firms in emerging economies as well. In addition to testing original and re-estimated versions of the models for the pre- and post-crisis periods, this research assesses 18 different holdout samples, which vary in the proportions of distressed and non-distressed populations in the original studies that proposed the models. A sensitivity analysis for each model also reflects the potential impact of country weights in the study sample. With this collected evidence, the current study suggests that researchers and practitioners can rely on the prediction models proposed by Ohlson (1980), Zmijewski (1984), and Shumway (2001) to decrease their decision-making costs. However, Altman's (1968) model cannot produce generalizable results for emerging economies, and Taffler's (1983) model requires caution, due to its mixed results.

To establish these contributions, this article starts with a review of literature pertaining to distress prediction models and their generalizability. Section 3 then outlines the data sources and sample description. After describing the methodology of the prediction models in Section 4, Section 5 details the results for the full and re-estimated samples, including holdout samples, pre- and post-financial crisis periods, and Type-I and Type-II errors obtained from sensitivity analyses. Finally, Section 6 concludes with some insights and limitations.

2. Literaturereview

Distress prediction studies respond to a practical need: to detect firm failure in advance (Ohlson, 1980). Early identification of bankruptcy or distress requires an accurate selection of financial ratios, appropriate estimation methods, and samples. Most researchers concentrate on developing new prediction models for new samples, starting with Altman's (1968) renowned *Z*-score model (e.g., Altman, Marco, & Varetto, 1994; Deakin, 1972; Gentry, Newbold, & Whitford, 1985b; Gilbert, Menon, & Schwartz, 1990; Lakshan & Wijekoon, 2013; Nam, Kim, Park, & Lee, 2008). They derive ratios and weightings from sample analyses, such that the results of the model tests are mostly sample specific. Yet as Mensah (1984) notes, the distribution of financial ratios changes over time, which implies that all models need to be reexamined periodically, in addition to being considered with a comparative view and with respect to their applicability to different economies.

2.1. Original models being tested

Altman (1968) offers the first distress prediction model, which emphasizes the classification performance achieved by financial ratios, in advance of an imminent failure. The model classifies 33 bankrupt and 33 non-bankrupt manufacturing firms during 1946 and 1965 with a multi-discriminant analysis estimation. Similar to the current study, its analysis of prediction ability relies on a continuing firm sample, and distress is defined as two or three years of negative income during the period 1958–1961.

Ohlson's (1980) model is derived from a sample of 105 bankrupt and 2058 non-bankrupt industrial firms traded on the stock exchange or over-the-counter market for the period between 1970 and 1976. The model's estimation method relies on a logistic regression.

Taffler (1983) examines the classification of 46 publicly listed, bankrupt manufacturing firms in the United Kingdom during 1969–1976. The estimation method applied for this classification process is linear discriminant analysis.

In Zmijewski's (1984) study, the classification ability of the proposed model is established with a sample of publicly listed industrial firms during 1972–1978. The sample consists of 81 bankrupt firms. To test prediction consistency, it considers different bankruptcy proportions from a matched sample of 40 bankrupt to 40 non-bankrupt firms, using a sample of 40 bankrupt and 800 non-bankrupt firms, together with a probit regression method.

Finally, Shumway (2001) proposes a prediction model on the basis of 300 publicly listed bankrupt firms and 2882 non-bankrupt industrial firms for the period between 1962 and 1992. It tests the impact of market variables on distress classifications and concludes that prediction models can improve their classification accuracies by incorporating accounting and market-based variables. This study uses a simple hazard model as its estimation method.

2.2. Developed economy implementations

An early study of the generalizability of prediction models by Begley, Ming, and Watts (1996) examines the applicability of Altman's (1968) and Ohlson's (1980) prediction models for U.S. publicly listed manufacturing firms, similar to the original studies, but during a more recent time period (1980–1989). With a sample of 165 bankrupt to 3300 non-bankrupt firms, they test both original and reestimated versions of each model and find that Ohlson's original model performs better than its reestimated form, whereas Altman's reestimated model provides better overall prediction accuracy than the original version. These results highlight the importance of checking prediction model generalizability, but Begley et al. only investigate U.S. samples, so the results are particular to the U.S. manufacturing industry.

Grice and Ingram (2001) also analyze the generalizability of Altman's (1968) prediction model, again with a U.S. industrial firm sample, for 1985–1991. They consider bankrupt, distressed, manufacturing, and non-manufacturing subsamples, with proportions that shift between 79 distressed to 452 non-distressed firms. They find that Altman's bankruptcy prediction model achieves accurate prediction for distressed and bankrupt samples; its reestimated form also improves classification accuracy. Similar to Begley et al. (1996), Grice and Ingram affirm the generalizability of Altman's model, to another U.S. sample, but they limit the number of observations to generalize the study results.

Then Grice and Dugan (2001) consider Ohlson's (1980) and Zmijewski's (1984) bankruptcy prediction models and their applicability for predicting both financial distress and bankruptcy. This study includes industrial and non-industrial U.S. firms that earn low stock and low bond ratings, over 1988 to 1999, and the subsamples include 183 distressed to 1043 non-distressed firms. In their results, Zmijewski's model provides better prediction accuracy for the distress classification, but for both models, classification performance has declined significantly compared with the original findings. Their study thus shows that these two prediction models are applicable to both bankruptcy and distressed samples, but it recommends that the Ohlson model should be used only for industrial firms, whereas the Zmijewski model has no industrial sensitivity. Here again though, the results are sample specific and restricted to the United States. In addition, Grice and Dugan do not reestimate the tested models, though they extended their analysis to such a reestimation subsequently (Grice & Dugan, 2003). Specifically, they reestimated the model coefficients for both the Ohlson and Zmijewski models. They do not specify the study period, but the observation numbers change from 181 distressed to 906 non-distressed firms across subsamples. Thus they show that the reestimated versions generate better classification accuracy than the originals for their U.S. samples.

Two other U.S.-based studies examine Altman's (1968) prediction model: Li (2012) and Li and Rahgozar (2012). They test its applicability for sample periods of 2008–2011 and 2000–2010, respectively, and concur that the original version of the model provides good prediction accuracy for bankrupt firms. However, the results again are sample specific and limited to bankrupt firms only.

A few studies apply bankruptcy prediction models to developed countries other than the United States. For example, with Canadian firm samples, Boritz et al. (2007) compare the bankruptcy prediction performance of Altman's (1968) and Ohlson's (1980) models against some Canadian models. Their study, spanning 1987–2002 and firm samples of 131 bankrupt and 135 non-bankrupt entries, identifies the Canadian prediction models as more accurate than either the Altman or Ohlson models. Yet the Altman and Ohlson models enhance bankruptcy detection capabilities, with their original coefficients. Boritz et al. (2007) thus test two prominent prediction models on a firm sample collected outside the United States, though their implementation still involves a single country.

Oude (2013) examines three models—Altman (1968), Ohlson (1980), and Zmijewski (1984)—with listed and non-listed Dutch firms. The sample includes 29 bankrupt and 802 non-bankrupt firms during 2008–2012. The results recommend reestimating these models for Dutch firms, due to the low classification accuracies achieved. However, these results also might not generalize, considering the relatively few firm observations and country-specific implications of each model.

2.3. Developing economy implementations

A few studies address less developed environments as well. For example, Sandin and Porporato (2007) scrutinize the applicability of Altman's (1968) model to listed Argentinian firms, namely, 11 bankrupt to 11 non-bankrupt firms during 1990–1998. Their findings suggest the applicability of the model for bankruptcy detection, but the small number of bankruptcy observations leaves the findings somewhat questionable.

Wang and Campbell (2010) test both original and reestimated versions of the Altman (1968) model for Chinese manufacturing firms, classifying delisted versus non-delisted firms for 2000–2008. With 42 delisted to 42 non-delisted firms, they find that Altman's model generates accurate classifications. But their results cannot signify whether the model also applies to other developing countries, because they are limited to the single-country setting of China and restricted in the number of firms.

Lawrence et al. (2015) turn their attention to Ohlson's (1980) model and its applicability to publicly listed firms in Thailand (60 distressed to 60 non-distressed). This prediction model was able to classify the distressed and non-distressed firms, yet this article did not reveal the study period or reestimate the model. In addition, it has a limited number of observations.

Finally, with a sample of Indian manufacturing firms, Singh and Mishra (2016) test the Altman (1968), Ohlson (1980), and Zmijewski (1984) models using two holdout samples: 130 bankrupt and 130 non-bankrupt, as well as 78 bankrupt and 78 non-bankrupt. The tests of original and reestimated versions for 2006–2014 indicate that using the reestimation enhances the prediction accuracy of each model for both holdout samples. That is, for each model, the original version produces mediocre classification accuracy, but its reestimated version significantly improves them. The study thus indicates the generalizability of the models through reestimation, though its results also might be sample specific.

As this literature review demonstrates, the bankruptcy prediction models have not been fully tested for developed and developing economies in combination. Despite the insights that prior studies offer, they cannot confirm the generalizability of the models for three main reasons. First, they each apply to a single country, such that the results are inherently sample specific. Second, the number of firm observations tend to be quite limited. Third, not every study includes reestimation efforts. To address these concerns, the current study tests the applicability of five bankruptcy prediction models on a cross-country basis, with sufficient firm observations. It also includes a reestimation of each prediction model (Begley et al., 1996; Grice & Dugan, 2003), to check the applicability to other samples that differ in their time periods, sampling criteria, and economic conditions.

3. Data

The data for this study come from 17 emerging market countries identified by Morgan Stanley Capital International (*MSC*I).[1] These data pertain to publicly listed industrial firms and the years between 2000 and 2012. The focus on publicly listed firms reflects the origins of all the prediction models, which were derived using information about publicly listed industrial firms in developed economies. With this study sample, this research should offer enhanced validity, because (1) the *MSC*I index integrates financial and accounting infrastructures in international markets. Thus, providing uniformity that improves the information quality of the variables, and (2) the stock prices and fundamental financial information for the emerging market countries have been gathered through the Bloomberg Professional and Thomson Reuters Eikon Data Terminals. Using these two databases maximizes the number of firm observations by extending the sample across countries.

Four additional criteria informed the sample selection. First, firm-year data had to come from fiscal year-end financial statements. Second, all firm-year observations must include each variable required by the five models tested in this study. Third, the firm-year data points must be continuous for the entire observation period. Fourth, outliers are removed at a 95% confidence level to improve the robustness of the study results. Therefore, the number of observations in the study sample decreased from 59,567 to 11,050.[2] The sector-specific sample breakdown identifies 34 industrial sectors in the country samples, as listed in Table 1.

This study reports separate descriptive statistics for overall, financially distressed versus healthy, and pre- and post-crisis samples. In Table 2, the variable distribution for the overall sample indicates that the observations are evenly distributed around the mean; the median values of the variables are close to the mean values for the full sample.

The distressed and non-distressed subsamples in Table 3 indicate significant differences across thegroups, especially in the performance indicators. The ratios are applicable to classify distressed and non-distressed firms with these estimation methods. For the current study, the dependent variable is two consecutive years with negative income. In this sense, the significance of the performance variables also appears important for predicting this dependent variable.

As a performance indicator, two consecutive years of negative income should be informative for any firm, but the financial crisis of 2008 also had deleterious effects on firm performance worldwide (Mihajlov, 2014; Notta & Vlachei, 2014; Oz & Balsari, 2017; Yap et al., 2014). Therefore, this study examines the models' responsiveness in terms of predicting performance distress during this catastrophic event and divides the sample into pre- and post-crisis subsamples, relative to the crisis year of 2008. Table 4 presents the descriptive results for these subsamples, which include 6800 and 4250 observations, respectively. Most of the variables differ significantly across the two subsamples; the impact of the crisis is evident in the averages achieved in each period. This natural test creates appropriate conditions for examining model prediction results under different economic conditions, for generalizability purposes.

4. Methodology

The current study scrutinizes the original and reestimated versions of the distress prediction models by Altman (1968), Ohlson (1980), Taffler (1983), Zmijewski (1984), and Shumway (2001) for a

[1] The countries are Brazil, Colombia, Egypt, Hungary, India, South Africa, Malaysia, Mexico, Morocco, the Philippines, Russia, South Korea, Taiwan, Thailand, Turkey, China, and Poland.

[2] The country-specific observation numbers are as follows: Brazil 65, Colombia 26, Egypt 182, Hungary 13, India 26, South Africa (Johannesburg) 286, Malaysia 13, Mexico 65, Morocco 156, Philippines 234, Russia 52, South Korea 2301, Taiwan 2301, Thailand 728, Turkey 585, China 3094, and Poland 923.

Table 1
Sample break down for each country and industry.

	Brazil	Colombia	Egypt	Hungary	India	South Africa	Malaysia	Mexico	Morocco	Philippines
Machinery, Tools, Heavy Vehicles, Trains & Ships			21.43%		50.00%	18.18%			25.00%	5.56%
Construction & Engineering		50.00%	50.00%		50.00%	31.82%		20.00%	41.67%	27.78%
Freight & Logistics Services			14.29%			9.09%			16.67%	11.11%
Professional & Commercial Services				100.00%		9.09%				5.56%
Transport Infrastructure	20.00%	50.00%	14.29%				100.00%	60.00%		16.67%
Electronic Equipment & Parts										
Passenger Transportation Services	40.00%					4.55%				11.11%
Diversified Industrial Goods Wholesalers						9.09%			8.33%	
Communications & Networking										
Metals & Mining						4.55%				
Aerospace & Defense	40.00%									
Industrial Conglomerates						4.55%		20.00%		5.56%
Investment Holding Companies						4.55%				
Real Estate Operations									8.33%	5.56%
Automobiles & Auto Parts										
Computers, Phones & Household Electronics										
Homebuilding & Construction Supplies										
Textiles & Apparel										
Media & Publishing										
Electric Utilities & IPPs										5.56%
Oil & Gas										5.56%
Construction Materials										
Food & Tobacco										
Pharmaceuticals										

	Russia	South Korea	Taiwan	Thailand	Turkey	China	Poland
Machinery, Tools, Heavy Vehicles, Trains & Ships		32.77%	49.15%	17.86%	26.67%	42.02%	26.76%
Construction & Engineering	25.00%	22.60%	9.60%	26.79%	13.33%	12.18%	32.39%
Freight & Logistics Services		19.21%	7.34%	5.36%	4.44%	4.62%	4.23%
Professional & Commercial Services		3.39%	2.82%	21.43%	20.00%	2.52%	16.90%
Transport Infrastructure	25.00%	1.13%	1.13%	8.93%	6.67%	11.34%	1.41%
Electronic Equipment & Parts		3.95%	13.56%			2.52%	
Passenger Transportation Services	25.00%	2.82%	1.69%	7.14%	2.22%	6.72%	
Diversified Industrial Goods Wholesalers		2.26%	0.56%		2.22%	5.46%	
Communications & Networking		2.26%	5.08%			0.84%	
Metals & Mining		2.82%	1.13%	1.79%	4.44%	0.84%	7.04%
Aerospace & Defense		1.69%			2.22%	2.94%	
Industrial Conglomerates	25.00%	1.13%		5.36%			
Investment Holding Companies		0.56%			6.67%	1.26%	
Real Estate Operations		0.56%	1.13%			1.26%	
Automobiles & Auto Parts		1.69%			2.22%	1.26%	
Computers, Phones & Household Electronics			2.82%				
Homebuilding & Construction Supplies		0.56%	0.56%		2.22%	0.42%	
Textiles & Apparel			0.56%		2.22%	0.84%	5.64%
Media & Publishing			0.56%	3.57%			
Electric Utilities & IPPs							1.41%
Oil & Gas						0.42%	1.41%
Construction Materials					2.22%		1.41%
Food & Tobacco						0.84%	
Pharmaceuticals						0.84%	

	Brazil	Colombia	Egypt	Hungary	India	South Africa	Malaysia	Mexico	Morocco	Philippines
Hotels & Entertainment Services						4.55%				
Diversified Retail										
Software & IT Services										
Semiconductors & Semiconductor Equipment										
Oil & Gas Related Equipment and Services										
Paper & Forest Products										
Food & Drug Retailing										
Specialty Retailers										
Healthcare Providers & Services										
Chemicals										
Country's total share	0.59%	0.24%	1.65%	0.12%	0.24%	2.59%	0.12%	0.59%	1.41%	2.12%

	Russia	South Korea	Taiwan	Thailand	Turkey	China	Poland
Hotels & Entertainment Services							
Diversified Retail			0.56%				

(*continued on next page*)

Table 1 (*continued*)

	Russia	South Korea	Taiwan	Thailand	Turkey	China	Poland
Software & IT Services			0.56%				
Semiconductors & Semiconductor Equipment			0.56%				
Oil & Gas Related Equipment and Services			0.56%				
Paper & Forest Products				1.79%			
Food & Drug Retailing					2.22%		
Specialty Retailers						0.42%	
Healthcare Providers & Services						0.42%	
Chemicals							1.41%
Country's total share	0.47%	20.82%	20.82%	6.59%	5.29%	28.00%	8.35%

Table 2
Descriptive statistics for overall sample.

	μ	$\tilde{\mu}$	σ	γ	κ	ε	N
CA/CL	2.99	1.38	35.17	35.89	1391.55	0.36	11,050
CA/TL	2.12	1.00	32.18	39.62	1679.06	0.31	11,050
CA-INV-CL/SALES-NIBT + DEPR	−2.69	0.31	103.71	2.65	645.62	0.99	11,050
CHIN	0.05	0.02	0.48	−0.14	0.42	0.00	11,050
CL/CA	0.92	0.72	1.69	23.88	840.38	0.02	11,050
EBIT/TA	0.06	0.05	0.35	95.11	9668.54	0.00	11,050
MVE/TL	7.66	2.12	13.81	3.27	12.54	0.15	11,050
OENEG	0.01	0.00	0.11	8.56	71.25	0.00	11,050
OCF/TL	9.65	5.21	41.46	−57.08	4779.52	0.40	11,050
PBT/ACL	0.71	0.12	49.72	104.86	1101.47	0.47	11,050
RE/TA	0.06	0.10	0.58	−16.26	378.05	0.01	11,050
Return	0.03	0.00	0.48	0.23	3.17	0.00	11,050
ROA	3.65	3.60	11.68	13.86	808.11	0.11	11,050
SALES/TA	0.98	0.76	4.64	60.69	4380.77	0.04	11,050
SIZE	3.06	2.87	0.73	2.06	8.72	0.01	11,050
SIGMA	0.03	0.03	0.03	10.04	151.95	0.00	11,050
TL/TA	0.54	0.53	0.34	11.82	307.63	0.00	11,050
WC/TA	0.12	0.13	0.33	−12.48	356.67	0.00	11,050

Notes: μ, $\tilde{\mu}$, σ, γ, κ, N and εrefers to Mean, Median, Standard Deviation, Skewness, Kurtosis, Number of observations and Standard Error respectively.

which represents the optimal point for a minimum cost classification. For samples established according to the original distress and non-distress proportion rates, the cut-off points reflect a scaling of the total distress observations to the total number in that specific observation sample.

A five-stage process supports the tests of the five models and their generalizability. The measure of generalizability is whether models produce high classification accuracies for each stage in emerging markets. First, the authors examine each model with the full study sample, to determine whether the original versions apply without any country specification. Their responsiveness, according to their original coefficients, to distress classifications would indicate the stability of the models. Second, this process continues by analyzing each reestimated version of the models with the full sample, to determine if their prediction accuracies improve or deteriorate for recent time periods. In this stage, the current study also measures the statistical significance of the reestimation results relative to the original model results.

Third, to explore the models' prediction function for the pre- and post-financial crisis, this study divides the sample at 2008. With this stage, the goal is to explore whether the predictions hold just prior to a financial crisis and after its universal effects on macro- and microeconomic factors. Fourth, this study establishes 18 holdout samples,

Table 3
Descriptive statistics for financially distressed and healthy firms.

	Financially distressed							Financially non-distressed							Mean dif.	
	μ	$\tilde{\mu}$	σ	γ	κ	ε	N	μ	$\tilde{\mu}$	σ	γ	κ	ε	N		
CA/CL	12.94	1.01	111.65	9.97	99.62	4.38	1099	2.29	1.41	20.66	75.79	6317.59	0.23	9101	* −5.622	(0.000)
CA/TL	7.57	0.64	85.48	13.02	170.65	2.58	1099	1.57	1.04	19.28	85.07	7648.60	0.20	9101	−0.027	(0.978)
CA-INV-CL/SALES-NIBT + DEPR	−1.81	−0.37	108.50	20.45	635.87	3.27	1099	−1.90	0.48	102.05	2.47	698.11	1.07	9101	*3.752	(0.000)
CHIN	0.00	0.00	0.37	0.07	0.92	0.01	1099	0.06	0.04	0.49	−0.17	0.35	0.01	9101	* −13.893	(0.000)
CL/CA	1.78	0.99	3.73	8.81	99.51	0.15	1099	0.85	0.71	1.35	34.42	1768.36	0.01	9101	*9.700	(0.000)
EBIT/TA	−0.04	−0.02	0.09	−1.85	6.34	0.00	1099	0.07	0.06	0.38	89.31	8333.29	0.00	9101	* −4.337	(0.000)
MVE/TL	5.56	1.22	10.85	3.95	20.66	0.38	1099	8.05	2.26	14.20	3.13	11.36	0.17	9101	* −16.228	(0.000)
OENEG	0.06	0.00	0.24	3.61	11.05	0.01	1099	0.01	0.00	0.08	12.52	154.76	0.00	9101	0.364	(0.716)
OCF/TL	−0.81	0.00	14.40	14.10	39.13	0.43	1099	11.50	8.27	45.00	−54.46	4202.83	0.48	9101	0.941	(0.347)
PBT/ACL	−0.62	−0.15	1.78	−6.41	53.60	0.05	1099	0.94	0.15	54.77	95.20	9075.15	0.57	9101	*22.148	(0.000)
RE/TA	−0.28	−0.06	0.87	−8.52	112.25	0.03	1099	0.11	0.12	0.50	−20.89	594.16	0.01	9101	*12.471	(0.000)
Return	−0.15	0.00	0.46	−0.27	7.05	0.01	1099	0.05	0.00	0.48	0.26	2.51	0.01	9101	*41.416	(0.000)
ROA	−8.61	−4.64	13.55	−3.29	22.49	0.41	1099	5.46	4.38	10.24	26.60	1626.79	0.11	9101	0.400	(0.689)
SALES/TA	0.93	0.56	5.68	18.70	352.67	0.17	1099	0.99	0.80	4.24	80.39	6990.62	0.04	9101	−0.188	(0.851)
SIZE	3.07	2.92	0.59	0.60	−0.17	0.02	1099	3.06	2.87	0.74	2.15	9.00	0.01	9101	* −11.818	(0.000)
SIGMA	0.04	0.03	0.04	7.40	71.25	0.00	1099	0.03	0.03	0.02	10.65	180.38	0.00	9101	* −9.441	(0.000)
TL/TA	0.63	0.57	0.65	8.19	110.29	0.02	1099	0.53	0.53	0.27	11.36	406.00	0.00	9101	*23.286	(0.000)
WC/TA	0.72	1.00	0.45	−0.98	−1.04	0.01	1099	0.57	1.00	0.50	−0.27	−1.93	0.01	9101	* −7.269	(0.000)

Notes: 1) *, ** and *** indicate statistical significance at the 1%, 5% and 10% level respectively. 2) μ, $\tilde{\mu}$, σ, γ, κ, N and ε refers to Mean, Median, Standard Deviation, Skewness, Kurtosis, Number of observations and Standard Error respectively.

multicountry sample of developing markets in a recent period. This application of the original and reestimated model coefficients includes the full, pre- and post-crisis, and holdout samples, and it checks the prediction results one year prior to firms' financial distress. For the optimal cut-off point for each matched sample, this study uses 0.5, similar to the matching and original sample proportions of each model. The reasons for defining these 18 holdout samples are dual: to strengthen the validity of the model results, and to compare the prediction results for small and large holdout samples with respect to the full sample. In turn, this study achieves a notably large holdout sample

Table 4
Descriptive statistics for pre -and post-crisis period.

	Pre-crisis sample							Post-crisis sample							Mean dif.	
	μ	$\tilde{\mu}$	σ	γ	κ	ε	N	μ	$\tilde{\mu}$	σ	γ	κ	ε	N		
CA/CL	3.75	1.37	46.82	27.27	794.05	0.64	6800	2.04	1.39	6.29	27.48	939.85	0.10	4250	**−2.359	(0.018)
CA/TL	2.64	0.98	40.96	31.18	1037.23	0.50	6800	1.29	1.03	2.69	44.99	2524.01	0.04	4250	**−2.158	(0.031)
CA-INV-CL/SALES-NIBT + DEPR	−4.24	0.14	125.73	2.56	467.61	1.52	6800	−0.20	0.58	51.61	−2.65	993.13	0.79	4250	**1.993	(0.046)
CHIN	0.09	0.02	0.45	−0.09	0.78	0.01	6800	−0.01	0.02	0.51	−0.12	0.00	0.01	4250	*−10.246	(0.000)
CL/CA	0.92	0.73	1.39	18.02	524.25	0.02	6800	0.92	0.72	2.00	24.72	804.89	0.03	4250	−0.015	(0.988)
EBIT/TA	0.06	0.05	0.44	76.96	6200.82	0.01	6800	0.06	0.05	0.08	0.56	6.70	0.00	4250	−0.543	(0.587)
MVE/TL	7.20	1.96	13.27	3.47	14.51	0.18	6800	8.39	2.40	14.60	2.99	10.12	0.25	4250	**−2.000	(0.046)
OENEG	0.02	0.00	0.12	7.82	59.21	0.00	6800	0.01	0.00	0.10	10.30	104.11	0.00	4250	*−2.882	(0.004)
OCF/TL	8.65	2.38	49.38	−54.79	3851.64	0.60	6800	11.26	8.64	23.68	−0.22	6.37	0.37	4250	−0.786	(0.432)
PBT/ACL	0.98	0.11	63.36	82.32	6784.78	0.77	6800	0.28	0.12	1.85	39.86	1992.70	0.03	4250	−0.719	(0.472)
RE/TA	0.04	0.08	0.52	−13.91	305.40	0.01	6800	0.10	0.13	0.66	−17.67	395.02	0.01	4250	*5.171	(0.000)
Return	0.05	0.00	0.44	0.68	5.49	0.01	6800	0.00	0.00	0.53	−0.11	0.98	0.01	4250	*−5.263	(0.000)
ROA	3.41	3.44	13.37	15.10	766.43	0.16	6800	4.04	3.83	8.30	−0.13	11.91	0.13	4250	*2.739	(0.006)
SALES/TA	1.04	0.75	5.88	48.34	2751.53	0.07	6800	0.90	0.78	0.76	4.79	41.87	0.01	4250	−1.479	(0.139)
SIZE	3.02	2.85	0.73	2.01	8.51	0.01	6800	3.12	2.90	0.72	2.20	9.26	0.01	4250	*7.112	(0.000)
SIGMA	0.03	0.03	0.03	9.00	123.25	0.00	6800	0.03	0.03	0.02	12.64	234.24	0.00	4250	*−6.306	(0.000)
TL/TA	0.52	0.51	0.36	10.84	244.87	0.00	6800	0.56	0.56	0.31	14.29	483.75	0.00	4250	*5.570	(0.000)
WC/TA	0.11	0.11	0.35	−11.14	271.05	0.00	6800	0.15	0.15	0.30	−15.64	582.71	0.00	4250	*7.142	(0.000)

Notes: 1) *, ** and *** indicate statistical significance at the 1%, 5% and 10% level respectively. 2) μ, $\tilde{\mu}$, σ, γ, κ, N and ε refers to Mean, Median, Standard Deviation, Skewness, Kurtosis, Number of observations and Standard Error respectively. 3) The number of financial distress observations are 791 to 308 respectively for pre-and post-crisis periods.

for distress observations. Fifth, the sensitivity analysis excludes some countries, then reapplies each model across the four aforementioned stages.

Specifically, this study applies the five-stage procedure to the following models:

Altman (1968):

$$Z = 0.012WC/TA + 0.014RE/TA + 0.033EBIT/TA + 0.006MVE/TL + 0.999SALES/TA$$

WC/TA:Working Capital to Total Assets

RE/TA:Retained Earnings to Total Assets

EBIT/TA:Earnings Before Interest and Taxes to Total Assets

MVE/TL:Market Value of Equity to Total Liability

SALES/TA:Sales to Total Assets

Ohlson (1980):

$$Z = -1.32 - 0.407SIZE + 6.03TL/TA - 1.43WC/TA + 0.0757CL/CA - 2.37NI/TA - 1.83OCF/TL - 1.72OENEG - 0.521CHIN + 0.285INTWO$$

SIZE = logarithm of total assets to GNP price − level index

TL/TA = Total Liabilities to Total Assets

WC/TA = Working Capital to Total Assets

CL/CA = Current Liabilities to Current Assets

NI/TA = Net Income to Total Assets

OCF/TL = Operational Cash Flows to Total Liabilities

OENEG = One if total liabilities exceeds total assets

CHIN = Change in Net Income

INTWO = One if net income was negative for the following two years , zero otherwise

Taffler (1983):

$$Z = 3.2 + 12.18PBT/ACL + 2.5CA/TL - 10.68CL/TA + 0.03(CA - INV - CL)/(SALES - NIBT + DEPR)$$

PBT/ACL = Profit Before Tax to Average Current Liabilities

CA/TL = Current Assets to Total Liabilities

CL/TA = Current Liabilities to Total Assets

INV = Inventory

NIBT = Net Income Before Tax

DEPR = Depreciation

Zmijewski (1984):

$$Z = -4.336 - 4.513NI/TA + 5.679TL/TA + 0.004CA/TL$$

NI/TA = Net Income to Total Assets

TL/TA = Total Liabilities to Total Assets

CA/TL = Current Assets to Total Liabilities

Shumway (2001):

$$Z = -13.30 - 0.48SIZE - 1.81RETURN - 1.98NI/TA + 5.79SIGMA + 3.59TL/TA$$

SIZE = Relative Size

RETURN = Annual Stock Return

NI/TA = Net Income to Total Assets

SIGMA = Standard Deviation of Daily Returns within a Year

TL/TA = Total Liabilities to Total Assets

5. Results

In presenting the comparative results, this study seeks to specify the level of generalizability of each prediction model (original and

Table 5
Original vs re-estimated coefficients.

Altman			Ohlson			Shumway			Zmijewski			Taffler		
X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$
WC/TA	0.012	*0.640	SIZE	−0.410	−0.009	SIZE	−0.480	***0.087	ROA	−4.530	*−0.093	PBT/ACL	12.18	**−0.100
RE/TA	0.014	we*0.563	TL/TA	6.030	*−0.316	RETURN	−1.810	*−0.295	TL/TA	5.679	−0.200	CA/TL	2.500	***0.026
EBIT/TA	0.033	*10.34	WC/TA	−1.430	*−0.423	ROA	−1.980	*−0.203	CA/TL	0.004	**0.003	CL/TA	−10.68	*2.385
MVE/TL	0.999	0.000	CL/CA	0.076	*0.004	SIGMA	5.790	**2.706				$\frac{CA-INV-CL}{SALES-NIBT-DEPR}$	0.030	−0.001
SALES/TA	0.006	***−0.058	ROA	−2.370	*−0.020	TL/TA	3.590	*−0.638						
			OCF/TA	1.830	*0.000									
			OENEG	−1.720	*0.636									
			CHIN	−0.521	**0.078									

Note: *, ** and *** indicate statistical significance at the 1%, 5% and 10% level respectively.

reestimated) to different samples. It also details the prediction accuracies of each model for the entire observation period, as well as the pre- and post-financial crisis periods. Furthermore, this section considers whether the models generate robust classification performance for different holdout samples and identifies the classification costs for each model according to the Type-I and Type-II error rates. Finally, to explore each model's sensitivity, this study reruns all the models after excluding some countries, according to the numbers of observations they represent in the study sample.

5.1. Original vs. reestimated coefficient and prediction results

The reestimated coefficients may be less stable, relative to the original coefficients, because the original coefficients were developed for different samples and in different time periods. The reestimated coefficients instead indicate the change in time, and the statistical significance of the variables should highlight the validity of the prediction models for different distressed samples. The reestimated coefficient significance also is indifferent to the original values for most variables. This situation might highlight similar characteristics for classifying distressed and non-distressed firms (Grice & Dugan, 2003). It also can establish whether the models are applicable for performance-based distress classification efforts (Table 5).

By running each model with the original coefficients and then comparing the prediction results with reestimated versions, this study attempts to determine whether any model(s) provide sustainable and generalizable classification outcomes.

5.1.1. Original model prediction results

Altman's (1968) model provides the least accurate classification for the full and the pre- and post-crisis samples. It is unresponsive to recent period distress classifications for emerging markets, with accuracy levels of 26.67% for the full and 30.22% and 23.11% for the pre- and post-crisis periods, respectively. Begley et al. (1996) and Grice and Ingram (2001) agree about the model's inability to develop predictions for recent periods in developed markets in the 1980s and 1990s. The Altman (1968) model is inappropriate for distress classifications in recent emerging markets.

The other four prediction models produce better distress classification accuracies for the full sample. Specifically, for the original model results, the Zmijewski, Taffler, Shumway, and Ohlson models have high classification accuracies, of 87.32%, 89.33%, 91.68%, and 94.32%, respectively, for the full sample results. In contrast, Grice and Dugan (2001) find better classification accuracy for the Zmijewski model during the 1980s and 1990s. In addition, the pre- and post-crisis prediction accuracies emphasize similar classification capabilities of all four models. The Zmijewski model provides the least accurate pre-crisis classification, at 83.98%; the Ohlson model provides the most accurate classification, at 92.47%. Their accuracy rankings are similar for the post-crisis period. The post-crisis classification performance of the Zmijewski model increases the most, possibly due to the adjustment of the inflated variables during the crisis period.

Furthermore, Taffler's (1983) model produces the third-highest classification accuracy for the full, pre-, and post-crisis periods, of 89.33%, 87.58%, and 91.08%, respectively. These results coincide with Agarwal and Taffler's (2007) finding that Taffler's original model provides high classification accuracy for a U.K. sample. Shumway's (2001) model achieves the second-highest classification ranking for the full, pre-, and post-crisis periods: 91.68%, 88.87%, and 94.50%, respectively. The market-based information (i.e., market size and stock returns) might encourage this incremental prediction performance, especially after the financial crisis, because these variables include more recent information than accounting-based variables.

5.1.2. Reestimated model prediction results

Reestimation is necessary to determine whether the original models are sufficiently responsive or if they require an update. Similar to the original model results, Altman's (1968) model produces the least accurate prediction after the reestimation of its coefficients, for the full, pre-, and post-crisis samples, reaching values of 27.56%, 33.78%, and 21.33%, respectively. The comparison of the original and reestimated prediction results from Altman's model for the emerging market sample is not statistically significant. Therefore, practitioners should be cautious in using this model for distress classification in emerging markets.

The incremental prediction accuracies of three reestimated versions, compared with their original versions, for the full sample are as follows: Taffler 93.56%, Zmijewski 94.14%, and Shumway 93.08%. That is, these three models should be subject to reestimations for emerging market samples. They yield high prediction accuracies with their original models, but the results indicate that their reestimated coefficients produce statistically significant, more accurate distress predictions.

The reestimated Ohlson (1980) model instead produces similar distress classification accuracy, 94.03%, to its original model, 94.32%, for the full sample. For the pre- and post-crisis periods, accuracy levels also are similar. However, the reestimation results reveal no statistical significance in the shift in prediction levels, according to binomial statistics. In addition to these binomial test statics, this study examines the significance of the results with a chi-square analysis, which supports the use of the Ohlson distress prediction model in emerging market samples with the original coefficients (Table 6).

In summary, this examination of prediction models for distress classification in emerging markets shows that the Zmijewski, Taffler, Shumway, and Ohlson models produce high prediction accuracies in both their original and reestimated versions. However, the statistical measures suggest that practitioners should reestimate the Zmijewski, Taffler, and Shumway models to improve classification accuracy; the Ohlson model instead is applicable with its original coefficients. In the subsample analyses, the Zmijewski, Taffler, and Shumway models produce statistically significant increases for the post-crisis period after the reestimation.

Table 6
Comparison of accuracy levels of each model for original and re-estimated models.

	Models				
	Altman	Zmijewski	Taffler	Shumway	Ohlson
Full sample					
Original	26.67%	87.32%	89.33%	91.68%	94.32%
Re-estimated	27.56%	93.56%	94.14%	93.08%	94.03%
Test statistics B	−1.488	*−15.77	*−12.98	*−3.922	0.920
Test statistics C	*3.923	*64.63	***1.959	*2.012	*16.79
Pre-crisis					
Original	30.22%	83.98%	87.58%	88.87%	92.47%
Re-estimated	33.78%	91.22%	92.24%	90.55%	92.04%
Test statistics B	*−5.673	*−16.32	*−11.50	*−4.110	1.196
Test statistics C	*2.918	*14.51	1.569	*1.722	*10.68
Post-crisis					
Original	23.11%	90.66%	91.08%	94.50%	96.18%
Re-estimated	21.33%	95.91%	96.05%	95.60%	96.02%
Test statistics B	*3.183	*−15.59	*−15.05	*−3.769	0.614
Test statistics C	*5.068	*102.3	*14.17	*2.577	*12.98

Notes: 1) *, ** and *** indicate statistical significance at the 1%, 5% and 10% respectively. 2) "*Test statistics B*" represents binomial tests and "*Test statistics C*" is the Chi-square test that compares the significance of the accuracy rate differences for original and re-estimated models.

5.2. Holdout sample results

The 18 holdout samples reflect each model's original distressed and non-distressed populations, which are important to consider as means to determine the models' performance for different samples. Reestimating each holdout sample for different numbers of firms also can reveal which models produce more consistent classification accuracies in terms of sample size. A holdout sampling process thus improves the validity of the results. Agarwal and Taffler (2007) note a small holdout sampling problem for non-distressed firms, and Grice and Ingram (2001) emphasize the size inefficiency of holdout samples. To address the size problem, the current study examines holdout samples composed of 50 to 1000 firms, reflecting the proportions of the original studies.

The results indicate that the Altman model is the least accurate distress classifier across the full, pre-, and post-crisis samples, through matched sampling. The reestimation results for the holdout samples indicate a slight improvement up to 300 firms, but the difference between the original and the reestimated model is insignificant. The Altman model again emerges as inappropriate for emerging markets.

The holdout sampling for the Zmijewski model reflects its original proportion of 1 distressed firm to 20 non-distressed ones. According to the prediction results for the full sample, this model provides a high level of prediction accuracy in its original version, and then the results significantly improve after reestimation. In addition, the significant increase in prediction accuracy persists even with larger samples. Therefore, practitioners should reestimate the Zmijewski model to support distress classifications in emerging markets.

Taffler's model relies on a matching process; its prediction accuracy decreases for holdout samples relative to the full sample. Although reestimation improves the results for the full sample, it diminishes the classification accuracy for the holdout samples. Therefore, the Taffler model should be exercised only with caution in emerging markets. Holdout sampling does not support high-level classification accuracy for either the original or reestimated versions of the model.

The Ohlson model generates a high level of accuracy for the full sample, with a ratio of 1 distressed to 20 non-distressed firms. However, the insignificance of the difference between the prediction results for the original versus reestimated versions affirms that this model is applicable with its original coefficients to emerging markets. There is no significant improvement in prediction results between the original and reestimated versions.

Table 7
Accuracy levels of each model for holdout samples.

Model	Holdout samples			
Altman	50 × 50	100 × 100	300 × 300	1000 × 1000
Original	39.24%	37.09%	30.02%	39.58%
Re-estimated	48.10%	43.71%	38.01%	36.85%
Test statistics B	−1.263	−1.349	*−2.921	***1.777
Test statistics C	*3.112	*5.023	**2.127	*3.204
Zmijewski	50 × 1000	100 × 2000	300 × 6000	
Original	92.00%	88.90%	87.98%	
Re-estimated	96.17%	96.85%	96.62%	
Test statistics B	*−4.050	*−10.01	*−18.19	
Test statistics C	*7.417	*10.95	*54.48	
Taffler	50 × 50	100 × 100	300 × 300	1000 × 1000
Original	73.42%	81.46%	73.43%	74.57%
Re-estimated	68.35%	67.55%	71.06%	70.98%
Test statistics B	0.789	*3.192	0.917	*2.550
Test statistics C	1.044	**2.033	0.166	***1.853
Shumway	50 × 450	100 × 900	300 × 2700	1000 × 9000
Original	94.46%	93.39%	93.08%	91.95%
Re-estimated	94.94%	94.59%	94.20%	93.52%
Test statistics B	−0.339	−1.129	***−1.777	*−4.277
Test statistics C	*21.00	*12.93	*77.12	*11.98
Ohlson	50 × 1000	100 × 2000	300 × 6000	
Original	97.14%	97.18%	97.45%	
Re-estimated	97.49%	97.47%	97.43%	
Test statistics B	−0.496	−0.582	0.071	
Test statistics C	*5.917	*15.86	*16.90	

Notes: 1) *, ** and *** indicate statistical significance at the 1%, 5% and 10% respectively. 2) "*Test statistics B*" represents binomial tests and "*Test statistics C*" is the Chi-square test that compares the significance of the accuracy rate differences for original and re-estimated models.

Finally, the Shumway model is based on 1 distressed and 9 non-distressed firms. The results of four holdout samples indicate that it produces accurate classification of distressed firms with both original and reestimated versions. However, sample size significantly affects their accuracy. In accordance with Shumway's (2001) initial study, which relies on a large sample for the distress classification, the current study results confirm better classification accuracy with larger sample sizes. The slight increase in prediction accuracy for the reestimated version, which is significant for large sample sizes, indicates that the Shumway model should be reestimated to improve the results in emerging markets (Table 7).

5.3. Type-I and Type-II error results

Using Type-I and Type-II errors, it is possible to specify the cost of mistakes in the Ohlson (1980), Taffler (1983), Zmijewski (1984), and Shumway (2001) models. Because the Altman (1968) model's prediction accuracy is very low, as discussed previously, it is not included in this error rate classification. A Type-I error entails classifying a financially distressed firm as non-distressed; a Type-II error is defined as the identification of a non-distressed firm as distressed.

The Type-I error rate is highest for the Zmijewski model and lowest for the Ohlson model, though the original versions of each model indicate that the latter invokes lower costs to identify financially distressed firms in emerging markets. The reestimation decreases the Type-I error rate for the Zmijewski, Shumway, and Taffler models, but it slightly increases this rate for the Ohlson model. However, as previously established, reestimation does not provide any significant change in the prediction results for the Ohlson model, so the change in error rates is insufficient to establish a valid interpretation for the model. The Type-II error rates are in general higher than the type-I error rates, specifically for the models' reestimation. However, lower type-I rates highlight the models' suitability for financial distress detection. For the Zmijewski, Shumway, and Taffler models though, it is necessary to update them to achieve better distress identification (Table 8).

Table 8
Type I and Type II results for original and re-estimated models.

Model	Original		Re-estimated	
	e_I	e_{II}	e_I	e_{II}
Ohlson	0.00%	6.46%	0.30%	6.37%
Shumway	3.71%	4.57%	1.95%	5.32%
Zmijewski	8.81%	2.79%	1.22%	5.79%
Taffler	5.72%	4.30%	0.14%	6.46%

Notes: 1) e_I, and e_{II} represents Type I and Type II errors respectively. 2) Each model uses original cut off points for prediction results. 3) Type I error:FD is classified as non-FD, Type II error: non-FD is classified as FD.

These error results support the findings that the model that provides the most accurate classification results in its original version is the Ohlson (1980) model. However, Zmijewski (1984), Taffler (1983), and Shumway (2001) enable improved prediction results with reestimation.

5.4. Model sensitivity results

A sensitivity analysis for each financial distress prediction model excludes China, Taiwan, and South Korea, due to their wide representation. That is, the total observations from these three countries represent approximately 70% of the study sample. Thus, the prediction models were reexamined to determine if they can achieve robust prediction accuracy for the remaining firm-year observations.

5.4.1. Original vs. reestimated coefficient and prediction results

After reestimating the coefficients, this study seeks to determine the generalizability of each prediction model with respect to its classification accuracy, assessed for the original and reestimated versions and for the pre- and post- crisis periods. The reestimation results highlight that the coefficients change, similar to the outcomes of the prediction models using the full sample. That is, the coefficients are unstable over time and vary with different samples, as might be expected. Furthermore, the reestimation results emphasize the statistically significant differences for most of the variables. Specifically, the reestimation of the models in the sensitivity analyses reveals that model coefficients change over time, so researchers and practitioners should examine both original and reestimated versions of the prediction models to determine whether they need to be updated (Table 9).

With regard to the original model results, the sensitivity analysis affirms that Altman's (1968) model produces the least accurate classification for pre-and post-crisis periods, similar to the overall sample. Again, Altman's model cannot predict financial distress for ongoing industrial firms in developing economies.

Zmijewski's (1984) model produces significant prediction accuracy before the sample firms suffer financial distress, and the pre-and post-crisis classification performance values support its robustness. That is, this model generates an equally solid distress classification for the subsamples as it does for the overall sample. Therefore, the original version of the model can be used.

The sensitivity results for Taffler's (1983) model also suggest its ability to distinguish financially distressed industrial firms from non-distressed firms. The accuracy levels for the entire observation period, as well as the pre-and post-crisis periods, reveal stable prediction performance by the model under different economic conditions. In addition, the sensitivity analysis produces accuracy levels that are similar to those for the full study sample, so the Taffler model can be applied to developing economies in its original form.

Shumway's (2001) model generates consistent classification performance for the sample periods as well. It offers an effective financial distress classifier, with its original coefficients and robust performance. The model produces similar predictions for the full sample results, emphasizing its applicability to developing economies.

In its original version, the Ohlson (1980) model is the most accurate financial distress prediction model throughout the observation period. This model sustains its solid classification performance for different economic conditions, such as in pre- and post-financial crisis periods. Thus, its original version can detect financial distress in advance for industrial firms in developing economies.

Overall, the sensitivity analysis for the prediction models suggests that the original version of each model achieves similar classification accuracy to that obtained for the total sample. The Altman model still is not proficient for classifying financially distressed firms in advance, whereas the other four prediction models offer full sample accuracy levels that range from 86.51% to 91.43%, highlighting their strength.

Turning to the reestimated model results, this study shows that the Altman (1968) model suffers reduced classification accuracy, beyond the already low level when using its original coefficients. Thus, the Altman distress model is not a good classifier in either its original or reestimated form for emerging economies.

For Zmijewski's (1984) model, the results suggest that the reestimation improves classification accuracy for financial distress prediction for each observation in the full, pre-, and post-crisis periods. These findings are supported by statistical test results at a 0.01 significance level. The sensitivity analysis thus supports the overall sample outcomes, with robust classification correctness. Even though Zmijewski's distress prediction model generates consistent classification accuracy with its original version in developing economies, practitioners and researchers might prefer a reestimation of this model, due to its promise of superior performance.

The Taffler (1983) model reestimation also reveals improved prediction accuracy for the full, pre-, and post-financial crisis samples. The distress classification indicates a statistically significant improvement at 0.01 in this sensitivity analysis, similar to the overall sample results. Thus, Taffler's model also should be subjected to reestimation, to improve on its already strong prediction accuracy in the original version, for emerging economies.

Table 9
Original vs re-estimated coefficients of each model.

Altman			Ohlson			Shumway			Zmijewski			Taffler		
X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$	X	β	$\widehat{\beta}$
WC/TA	0.012	* −3.951	SIZE	−0.410	* −0.315	SIZE	−0.480	* −0.436	ROA	−4.530	* −0.188	PBT/ACL	12.18	* −5.933
RE/TA	0.014	−0.028	TL/TA	6.030	0.399	RETURN	−1.810	** −0.239	TL/TA	5.679	*1.357	CA/TL	2.500	0.005
EBIT/TA	0.033	* −26.18	WC/TA	−1.430	** −0.842	ROA	−1.980	* −0.196	CA/TL	0.004	**0.008	CL/TA	−10.68	*0.002
MVE/TL	0.999	* −0.009	CL/CA	0.076	0.039	SIGMA	5.790	−1.279				$\frac{\text{CA} - \text{INV} - \text{CL}}{\text{SALES} - \text{NIBT} - \text{DEPR}}$	0.030	0.000
SALES/TA	0.006	* −0.170	ROA	−2.370	* −0.212	TL/TA	3.590	*** −0.311						
			OCF/TA	1.830	−0.000									
			OENEG	−1.720	0.301									
			CHIN	−0.521	*0.795									

Notes: 1) The table results represent the sample after the exclusion of China, Taiwan and South Korea. 2) X represents the independent variables of each model, β stands for the original coefficients, and $\widehat{\beta}$ indicates re-estimated coefficients. 3) *, ** and *** indicate statistical significance at the 1%, 5% and 10% level respectively.

Table 10
Accuracy levels of each model for full sample and pre- & post-crisis periods.

	Model				
	Altman	Zmijewski	Taffler	Shumway	Ohlson
Full sample					
Original	22.91%	86.51%	87.51%	89.76%	91.43%
Re-estimated	9.18%	90.85%	90.96%	90.35%	90.35%
Test statistics B	*1.560	*2.981	*1.946	*2.010	*1.374
Test statistics C	*9.627	*13.37	*1.369	*77.86	*12.46
Pre-crisis					
Original	25.19%	83.41%	84.34%	85.57%	88.60%
Re-estimated	12.48%	86.89%	87.98%	86.21%	86.21%
Test statistics B	*1.373	*2.430	*1.594	*1.732	*2.238
Test statistics C	*5.276	*5.363	*1.201	*14.89	*9.623
Post-crisis					
Original	20.62%	89.61%	90.69%	93.94%	94.26%
Re-estimated	5.89%	94.80%	93.95%	94.50%	94.50%
Test statistics B	*1.969	*4.514	*2.391	*2.647	*3.462
Test statistics C	*9.903	*20.22	*1.369	*39.37	*9.321

Notes: 1) The table results represent the sample after the exclusion of China, Taiwan and South Korea. 2) *, ** and *** indicate statistical significance at the 1%, 5% and 10% respectively. 3) "*Test statistics B*" represents binomial tests and "*Test statistics C*" is the Chi-square test that compares the significance of the accuracy rate differences for original and re-estimated models.

Table 11
Accuracy levels of each model for holdout samples.

Model	Holdout samples		
Altman	50 × 50	100 × 100	300 × 300
Original	40.51%	37.74%	39.95%
Re-estimated	21.51%	14.57%	11.85%
Test statistics B	1.177	1.204	*1.980
Test statistics C	*2.676	*3.414	*7.226
Zmijewski	50 × 1000	100 × 2000	
Original	83.44%	84.01%	
Re-estimated	95.55%	95.05%	
Test statistics B	*1.547	*1.159	
Test statistics C	*13.61	*11.70	
Taffler	50 × 50	100 × 100	300 × 300
Original	77.33%	72.48%	74.01%
Re-estimated	66.66%	69.12%	67.24%
Test statistics B	*17.03	*34.60	*55.31
Test statistics C	1.050	**1.979	*5.273
Shumway	50 × 450	100 × 900	300 × 2700
Original	91.19%	91.60%	90.43%
Re-estimated	91.42%	92.55%	91.67%
Test statistics B	*1.368	*1.485	*1.779
Test statistics C	*20.50	*277.5	*89.53
Ohlson	50 × 1000	100 × 2000	
Original	97.22%	97.44%	
Re-estimated	96.66%	96.94%	
Test statistics B	*5.807	*6.462	
Test statistics C	*11.15	*9.597	

Notes: 1) The table results represent the sample after the exclusion of China, Taiwan and South Korea. 2) *, ** and *** indicate statistical significance at the 1%, 5% and 10% respectively. 3) "*Test statistics B*" represents binomial tests and "*Test statistics C*" is the Chi-square test that compares the significance of the accuracy rate differences for original and re-estimated models.

The reestimation of Shumway's (2001) prediction model generates a slight enhancement of the classification results. The prediction results from each observation period indicate improvements and support the overall sample findings, which suggest the benefits of reestimating this model. Specifically, both binomial and chi-square tests achieve significance at the 0.01 level, indicating the need for reestimation to achieve superior distress classification.

Only the Ohlson (1980) model reestimation does not suggest updating the model coefficients. Rather, the model generates superior distress classification in its original version for developing economies. This finding from the sensitivity analysis is in line with the overall sample results; the statistical test results also indicate the significance of the original model (Table 10).

In summary, the sensitivity analyses for original and reestimated distress models produce results similar to those for the overall study sample. That is, the Altman model still does not achieve successful classifications, in its original or reestimated versions, but the other four models perform better. The analyses suggest reestimating the Zmijewski, Taffler, and Shumway models to produce even better predictions. The average full sample reestimated prediction accuracy of these latter distress prediction models is approximately 91%.

5.4.2. Holdout sample results

The sensitivity results support the overall sample findings, but another test of the generalizability of the prediction models relies on different sample sizes, to identify any potential size effect on prediction accuracy (Grice & Ingram, 2001). Therefore, this study establishes 13 holdout samples to assess each prediction model's original distressed to non-distressed population proportions.

The Altman (1968) model results indicate the lowest classification accuracy, for both its original and reestimated versions, across three different matched holdout samples, as in the original study. These results confirm the low overall sample performance of the model. The Altman model generates the most inferior prediction accuracy and should not be used for distress classifications.

For the Zmijewski (1984) model, the results reveal improved classification accuracy through reestimation, regardless of the differences in the holdout sample sizes. Therefore, the Zmijewski model should be reestimated for developing country samples to improve its distress classification accuracy. Similarly, the holdout sample results for the Shumway (2001) model suggest reestimation to slightly improve its classification accuracy. This improvement is particularly evident with larger sample sizes with respect to the distressed and non-distressed proportions. The original version of this model generates strong classification accuracy, but the reestimation results are better, as indicated by the statistical significance of the prediction results.

The outcome of the Taffler (1983) prediction model highlights the stability of its original coefficients for distress prediction. The original and reestimated coefficient results for three different matched holdout samples suggest implementing the model in its original version. For the Ohlson (1980) model, prediction accuracy also is better in its original version. The model generates approximate classification accuracies for two holdout samples with different sizes using its original coefficients, but the reestimation produces inferior performance for financial distress prediction (Table 11).

The sensitivity analysis with various holdout samples reinforces the overall sample results. That is, the Altman (1968) model does not yield strong classification accuracy, in its original or reestimated version, in developing economies. The Zmijewski (1984) model generates successful distress predictions in its original and reestimated versions, but the reestimation results significantly surpass the original version. The Taffler (1983) model offers a reasonably correct classification for financially distressed firms using the original coefficients, and its reestimation reduces the accuracy of the classification performance. The Shumway (2001) model is one of the best distress classifiers, in both its original and reestimated forms. The statistical significance outcomes suggest that reestimation reinforces the correct classification of distressed firms. The Ohlson (1980) model is the best classifier of financial distress with its original coefficients. The robustness of the original model results is supported by the statistical tests of the current study and in line with findings by Begley et al. (1996) and Boritz et al. (2007).

5.4.3. Type-I and Type-II error results

Similar to the overall sample results, the Altman model is excluded from this sensitivity analysis of the error rates. The findings suggest that the Ohlson model, which has the lowest Type-I error rate, is the most

Table 12
Type I and Type II results for original and re-estimated models.

Model	Original		Re-estimated	
	e_I	e_{II}	e_I	e_{II}
Ohlson	0.00%	8.62%	0.30%	8.35%
Shumway	4.20%	5.31%	2.26%	6.79%
Zmijewski	9.60%	2.62%	1.52%	7.48%
Taffler	6.11%	5.21%	0.35%	8.61%

Notes: 1) The table results represent the sample after the exclusion of China, Taiwan and South Korea. 2) e_I, and e_{II} represents Type I and Type II errors respectively. 3) Type I is the selection of a financially distressed firm as non-distressed and Type II is the selection of financially non-distressed firm as distressed.

cost efficient when using its original coefficients. That is, the original Ohlson model offers the most proficient distress prediction in terms of detecting financially distressed firms in advance. The reestimation of this model reduces its classification accuracy. The Shumway model has the second-lowest Type-I error rate with its original coefficients, and reestimation improves these results, in line with the overall sample error rates. The Zmijewski model's Type-I error rate is highest before reestimation, but the update significantly improves its distress classification accuracy. For the Taffler model, the original version produces the third-lowest Type-I error rate, and the reestimation improves its distress detection capability and decreases the cost of misdetection for financially distressed firms (Table 12).

A low Type-I error rate is necessary to consider a prediction model a satisfactory distress classifier. The findings highlight that the Ohlson model is the best distress predictor with its original coefficients; the Shumway, Zmijewski, and Taffler models need to be reestimated to reduce Type-I errors.

6. Conclusion

This study contributes to extant literature by confirming the level of generalizability of some prominent prediction models. Some of these models produce successful distress predictions across different samples, times, and economic conditions. Therefore, they can serve as benchmarks for continued studies in emerging markets, enabling researchers to continue improving existing prediction models or establish new ones. In particular, the results of the current study confirm that, other than the Altman (1968) model, popular distress prediction models are generalizable to different samples and time periods. According to a detailed, five-stage examination process, practitioners in emerging markets should reestimate the Zmijewski (1984) and Shumway (2001) prediction models, because doing so leads to statistically significant improvements in all prediction tests (i.e., full and subsample outcomes and holdout sampling). In contrast, the Ohlson (1980) model provides high prediction accuracy in its original version, indicating that it is stationary in time and applicable to distinct time periods and economic conditions. Despite its high prediction accuracy for the full sample and the pre- and post- financial crisis periods, the holdout sample results for the Taffler (1983) model do not support improved prediction processes though, so researchers should use caution before applying this model to developing country samples.

The current study results pertain to a broad range of *MSC*I emerging market countries, yet even in this case, the various economies feature specific financial and accounting infrastructures, relative to those in place in other developing countries. The results thus are valid for the publicly listed firms in these samples, and they should be examined for non-listed firms. Further studies could also examine more publicly listed firms in different developing countries, depending on the availability of the data needed to test the models' generalizability for additional samples.

Conflict of interest

No conflicts of interest are declared.

References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance, 32*(8), 1541–1551.
Agarwal, V., & Taffler, R. J. (2007). Twenty-five years of the Taffler z-score model: Does it really have predictive ability? *Accounting and Business Research, 37*(4), 285–300.
Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance, 36*, 278–285.
Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.
Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, W. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting, 28*(2), 131–171.
Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance, 18*(3), 505–529.
Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies, 1*(4), 267–284.
Boritz, J. E., Kennedy, D. B., & Sun, J. Y. (2007). Predicting business failures in Canada. *Accounting Perspectives, 6*(2), 141–165.
Charitou, A., Neophytou, E., & Charalambous, C. (2004). Predicting corporate failure: Empirical evidence for the UK. *The European Accounting Review, 13*(3), 465–497.
Chouhan, V., Chandra, B., & Gosvami, S. (2014). Predicting financial stability of select BSE companies revisiting Altman Z-score. *International Letters of Social and Humanistic Sciences, 26*, 92–105.
Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research, 10*(1), 167–179.
DeAngelo, H., & DeAngelo, L. (1990). Dividend policy and financial distress: An empirical investigation of troubled NYSE firms. *Journal of Finance, 45*(5), 1415–1431.
Gentry, J. A., Newbold, P., & Whitford, D. T. (1985a). Classifying bankrupt firms with funds flow components. *Journal of Accounting Research, 23*(1), 146–160.
Gentry, J. A., Newbold, P., & Whitford, D. T. (1985b). Predicting bankruptcy: If cash flow's not the bottom line, what is? *Financial Analysts Journal, 41*(5), 47–56.
Gilbert, L. R., Menon, K., & Schwartz, K. B. (1990). Predicting bankruptcy for firms in financial distress. *Journal of Business, Finance & Accounting, 17*(1), 161–171.
Grice, J. S., & Dugan, M. T. (2001). The limitations of bankruptcy prediction models: Some cautions for the researcher. *Review of Quantitative Finance and Accounting, 17*, 151–166.
Grice, J. S., & Dugan, M. T. (2003). Re-estimations of the Zmijewski and Ohlson bankruptcy prediction models. *Advances in Accounting, 20*, 77–93.
Grice, J. S., & Ingram, R. W. (2001). Tests of generalizability of Altman's bankruptcy prediction models. *Journal of Business Research, 54*, 53–61.
Hill, N. T., Perry, S. E., & Andes, S. (1996). Evaluating firms in financial distress: An event history analysis. *Journal of Applied Business Research, 12*(3), 60–71.
Hussain, F., Ali, I., Ullah, S., & Ali, M. (2014). Can Altman Z-score model predict business failures in Pakistan? Evidence from textile companies of Pakistan. Journal of Economics and. *Sustainable Development, 5*(13), 110–115.
Khanna, T., & Palepu, K. G. (2006). Emerging giants: Building world-class companies in developing countries. *Harvard Business Review, 84*(10).
Kordlar, A. E., & Nikbakht, N. (2011). Comparing bankruptcy prediction models in Iran. *Business Intelligence Journal, 4*(2), 335–342.
Lakshan, A. M. I., & Wijekoon, W. M. H. N. (2013). The use of financial ratios in predicting corporate failure in Sri Lanka. *GSTF Journal of Business Review (GBR), 2*(4), 37–43.
Lawrence, J. R., Pongstat, S., & Lawrence, H. (2015). The use of Ohlson's O-Score for bankruptcy prediction in Thailand. *Journal of Applied Business Research, 31*(6), 2069–2078.
Li, H., & Sun, J. (2008). Ranking-order case-based reasoning for financial distress prediction. *Knowledge-Based Systems, 21*, 868–878.
Li, J. (2012). Prediction of corporate bankruptcy from 2008 through 2011. *Journal of Accounting and Finance, 12*(1), 31–41.
Li, J., & Rahgozar, R. (2012). Application of the Z-score model with consideration of total assets volatility in predicting corporate financial failures from 2000-2010. *Journal of Accounting and Finance, 12*(2), 11–19.
Lifschutz, S., & Jacobi, A. (2010). Predicting bankruptcy: Evidence from Israel. *International Journal of Business and Management, 5*(4), 133–141.
Malik, U. S., Aftab, M., & Noreen, U. (2013). Distress risk and stock returns in an emerging market. *Research Journal of Finance and Accounting, 4*(17), 81–85.
Mensah, Y. M. (1984). An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. *Journal of Accounting Research, 22*(1), 380–395.
Mihajlov, D. K. (2014). Profitability during the financial crisis: Evidence from the regulated capital market in Serbia. *South-Eastern Europe Journal of Economics, 1*, 7–33.
Mizan, A., & Hossain, M. (2014). Financial soundness of cement industry of Bangladesh: An empirical investigation using Z-score. *American Journal of Trade and Policy, 1*(1), 16–22.

Nam, C. V., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting, 27*(6), 493–506.

Notta, O., & Vlachei, A. (2014). The impact of financial crisis on firm performance in case of Greek food manufacturing firms. *Procedia Economics and Finance, 14*, 454–460.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

Oude, A. J. (2013). *Testing the generalizability of the bankruptcy prediction models of Altman, Ohlson and Zmijewski for Dutch listed and large non-listed firms.* ThesisUniversity of Twente1–61.

Oz, I. O., & Balsari, K. C. (2017). Growth strategy of a contract automobile manufacturer in Turkey during 2008-2009 financial crisis. *Asian Journal of Management Cases, 14*(1), 49–64.

Oz, I. O., & Yelkenci, T. (2015). The generalizability of financial distress prediction models: Evidence from Turkey. *Accounting and Management Information Systems, 14*(4), 685–703.

Oz, I. O., & Yelkenci, T. (2017). A theoretical approach to financial distress prediction modeling. *Managerial Finance, 43*(2), 212–230.

Pongsatat, S., Ramage, J., & Lawrence, H. (2004). Bankruptcy prediction for large and small firms in Asia: A comparison of Ohlson and Altman. *Journal of Accounting and Corporate Governance, 1*(2), 1–13.

Sandin, A. R., & Porporato, M. (2007). Corporate bankruptcy prediction models applied to emerging economies: Evidence from Argentina in the years 1991–1998. *International Journal of Commerce and Management, 17*(4), 295–311.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74*(1), 101–124.

Singh, B. P., & Mishra, A. K. (2016). Re-estimation and comparisons of alternative accounting based bankruptcy prediction models for Indian companies. *Financial Innovation, 2*(6), 1–28.

Taffler, R. J. (1983). The assessment of company solvency and performance using a statistical model. *Accounting and Business Research, 13*(52), 295–307.

Wang, Y., & Campbell, M. (2010). Business failure prediction for publicly listed companies in China. *Journal of Business and Management, 16*(1), 75–88.

Yap, B. C.-F., Mohamed, Z., & Chong, K. R. (2014). The effects of the financial crisis on the financial performance of Malaysian companies. *Asian Journal of Finance & Accounting, 6*(1), 236–248.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–82.