

Predicting Corporate Bankruptcy Risk in Australia: A Latent Class Analysis

Stewart Jones*
Maurice Peat*

Abstract

In this paper a latent class model (LCM) is applied to estimate corporate bankruptcy and insolvency risk in Australia using a number of financial, market and macro-economic variables and indicators. LCMs represent a significant improvement on traditional techniques such as standard logit and linear discriminant analysis because they relax the highly restrictive IID condition which can distort parameter estimates and potentially undermine predictive accuracy (Jones and Hensher, 2004).

While LCMs are more general (powerful) than standard approaches, they differ from many other non-IID approaches in that they are relatively straight forward to estimate and interpret. In this study we demonstrate the application and interpretation of LCM models based on a large sample of corporate failures in Australia. We also consider the potential of LCMs for future research and practice in this field.

Keywords

Latent Class Model (LCM)
Logit Analysis
Discriminant Analysis
Bankruptcy in Australia
Insolvency Risk
IID Condition

Introduction

The bankruptcy modeling literature has moved to more sophisticated discrete choice models which relax the highly restrictive IID condition (*independent and identically distributed errors*) and the IIA assumption (*independence of irrelevant alternatives*) inherent in standard logit and probit models (Jones and Hensher, 2004). While advanced logit models, such as mixed logit, are more difficult to estimate and interpret, they have a number appealing statistical properties and are better equipped to handle firm specific heterogeneity (both within and across firms) which can distort parameter estimates and undermine predictive performance. Jones and Hensher (2007) extended their original study on mixed logit models to consider other advanced logit structures, including nested logit. They find that nested logit also outperform standard logit models by significant margins, and overall model performance is comparable to a mixed logit model.

However, the LCM is an advanced logit structure that is considerably easier to estimate and interpret than mixed logit as the model has a closed form solution. In fact, the LCM proposed here is essentially a semi-parametric variant of the mixed logit model (see Train, 2003).¹ The LCM can also handle the highly restrictive IID and IIA conditions, and unlike the mixed logit model, provides a globally optimal set of parameter estimates. A particular limitation of the mixed logit model is the relative complexity in estimation. For instance, the mixed logit model has an open form solution and thus requires analytical integration and use of simulated maximum likelihood to estimate model parameters and changes in outcome probabilities. Mixed logit models also do not provide a single set of globally optimal parameter estimates (i.e., due to the requirement for simulated maximum likelihood) and assumptions must be imposed for the distribution of unobserved influences (Jones and Hensher 2007).

¹For instance, for a mixed logit model, the mixing distribution $f(\beta)$ is assumed to be a continuous variable (see Jones and Hensher, 2004). However, if we assume β takes on a finite distinct set of values (ie. The distribution of β is discrete), we have in effect a latent class model (the formal proof is provided in Train, 2003, pp.139-140).

* University of Sydney

A simple illustration is proposed by Goodman (2002) and discussed in Jones and Hensher (2008). Consider the simplest of cases of a cross-classification of analysis of two dichotomous variables which has a two way 2 x 2 cross classification table [X ,Y]; where the two rows of the 2 x 2 table correspond to the two classes of the dichotomous variable X, and the two columns of the 2 x 2 table correspond to the two classes of the dichotomous variable Y. Let P_{ij} denote the probability that an observation will fall in the i th row ($i = 1,2$) and j th column ($j = 1,2$) of this 2 x 2 table. If the variables X and Y are statistically independent, we have the following simple relationship (ie., the assumption of local independence):

$$P_{ij} = P_i^X P_j^Y \tag{1}$$

Where P_i^X is the probability that an observation will fall in the i th class on variable X (the i th row of the 2 x 2 table), and P_j^Y is the probability that an observation will fall in the j th class (the j th column of the 2 x 2 table) on variable Y with

$$P_i^X = P_{i+} = \sum_j P_{ij}, P_j^Y = P_{+j} = \sum_i P_{ij} \tag{2}$$

A practical application of this simple concept is provided by Lazarsfeld and Henry (1968). Suppose that a sample of 1000 people are asked whether they read journal X and Y with the survey responses appearing as follows:

	Read X	Did not read X	Total
Read Y	260	140	400
Did not read Y	240	360	600
Total	500	500	1000

It can be readily see that the two variables (reading X and reading Y) are strongly related (the chi square test is statistically significant), and therefore X and Y are not independent of each other. Readers of X tend to read Y more often (52%) than non readers of X (28%).

When reading X and Y is independent, than $P(X\&Y) = P(X)*P(Y)$. However, $260/1000$ is not $400/1000*500/1000$. Thus reading X and Y is dependent of each other. However, adding the education level of respondents generates the following table:

High education	Read X	Did not read X	Total	Low education	Read X	Did not read X	Total
Read Y	240	60	300	Read Y	20	80	100
Did not read Y	160	40	200	Did not read Y	80	320	400
Total	400	100	500	Total	100	400	500

And again if reading X and Y are independent, than $P(A\&B) = P(A)*P(B)$ for each education level.

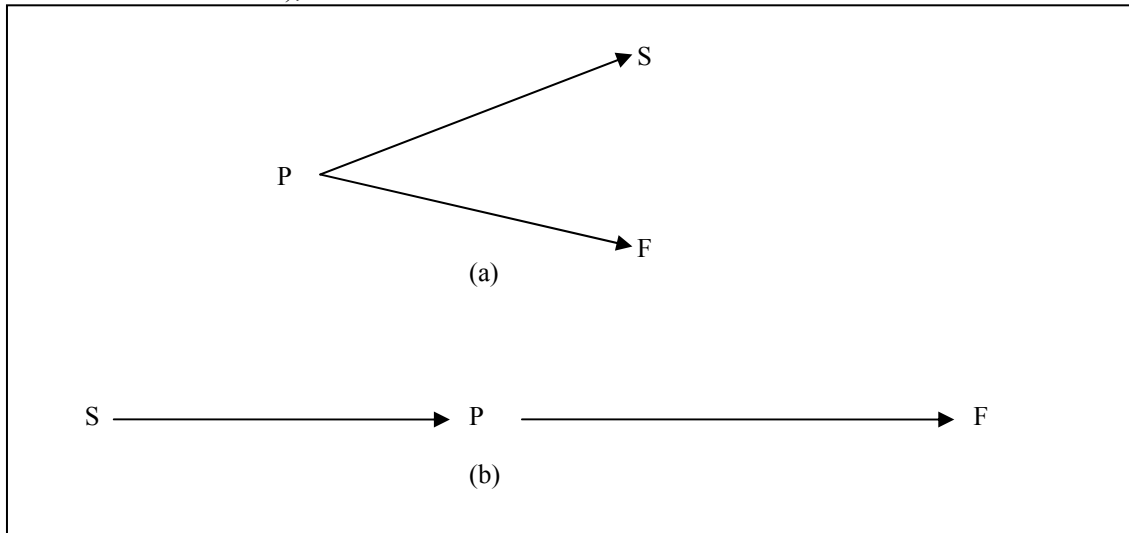
Note that:
 $240/500 = 300/500* 400/500$ and $20/500 = 100/500*100/500$

Hence, when we examine separately the high and low educated people, there is no relationship between the two journals (i.e., reading X and Y are independent within educational level). The educational level accounts for the difference in reading X and Y. When variables X and Y are not statistically independent, (1) does not hold. If X and Y are

key variables of interest to the analyst, the analyst would be interested in measuring the degree of non-interdependence (or correlation) between X and Y. While there are many measures of association and correlation that can reveal the magnitude of non interdependence between X and Y, Jones and Hensher (2008) point out they cannot determine whether the relationship between X and Y is spurious; that is whether the apparent relationship between X and Y can be explained away (or even explained more fully) by some other variable, say Z, where this variable may be unobserved or latent.

A simple illustration with firm failures is provided by Jones and Hensher (2008). A statistically significant relationship between firm size (S) (measured by market capitalization) and corporate failure (F) is often observed in this research (ie., smaller public companies on average tend to have a higher propensity to fail than larger public companies). However, it is possible that any number of latent effects or factors could influence this relationship. Let us consider one such factor, which we call firm financial performance (P). It is possible that P could be driving both S and F (so P is an antecedent variable to both S and F), in which case S and

F are conditionally independent of each other given the level of P (see Figure 1 (a) below) (see Goodman, 2002). That is, higher performing companies tend to be associated with higher stock prices and therefore higher market capitalizations (ie., these firm are larger on average); furthermore firms with better overall financial performance tend to have a lower probability of failure relative to poorer performing firms). Hence, the apparent relationship between S and F could be spurious.



Another possible scenario is that firm size (S) could also be driving financial performance (P) which in turn drives F, in which case P is an intervening variable as shown in Figure 1 (b). In this case, larger firms tend to have higher market concentrations, greater access to capital and consumer markets and greater economies of scale in production which could lead to superior overall financial performance. Again, as noted by Jones and Hensher (2008), S and B are conditionally independent, given the level of P. It is also possible that S and P could also be reciprocally affecting each other relation where S drives P. Again, S and F are conditionally independent.

More formally, the LCM assumes that firms are implicitly sorted into a set of Q latent classes, but which latent class contains any particular firm is unknown to the researcher. When the dependent variable is ordinal or nominal, the central behavioral LCM model is a logit model for discrete outcomes among J_i

alternatives, by firm i observed in T_i outcome situations,

$$P_{jit|q} = \frac{\exp(X'_{itj}\beta_q)}{\sum_{j=1}^J \exp(X'_{itj}\beta_q)} \quad (3)$$

The latent class q assignment is unknown. Let H_{iq} denote the prior probability for latent class q for firm i . Various formulations have been used this (see Greene 2003) but the most common form is set out as follows:

$$H_{iq} = \frac{\exp(z'_i\theta_q)}{\sum_{q=1}^Q \exp(z'_i\theta_q)} \quad (4)$$

where z_i denotes a set of observable characteristics which enter the model for class membership and θ_q denotes the latent class parameter vectors.

Determining the Number of Latent Classes

An issue to be confronted is the choice of Q , the number of latent classes. As pointed out in Jones and Hensher (2008), if there is a known Q^* that is greater than the ‘true’ Q , then it is possible to ‘test down’ to Q by using, for example likelihood ratio tests. A model with $Q+1$ classes encompasses one with Q if the parameters in any two of the $Q+1$ classes are forced to equality. This does move the problem up one level, since the Q^* must now be assumed known, but testing down from a specified Q^* is straightforward. (‘Testing up’ from a small Q (one) is not valid, since the estimates obtained for any model that is too small are inconsistent.) Roeder et al. (1999) suggest using the Bayesian Information Criterion or BIC:

$$BIC(\text{model}) = \ln L + \frac{(\text{model size}) \ln N}{N} \quad (5)$$

With the parameter estimates of θ_q in hand, the prior estimates of the class probabilities are \hat{H}_{iq} . Using Bayes theorem, we can obtain a posterior estimate of the latent class probabilities using

$$\hat{H}_{q|i} = \frac{\hat{p}_{i|q} \hat{H}_{iq}}{\sum_{q=1}^Q \hat{p}_{i|q} \hat{H}_{iq}} \quad (6)$$

As described in Jones and Hensher (2008), the notation $\hat{H}_{q|i}$ is used to indicate the firm-specific estimate of the class probability, conditioned on their estimated outcome probabilities, as distinct from the unconditional class probabilities which enter the log likelihood function. A strictly empirical estimator of the latent class within which the individual resides would be that associated with the maximum value of $\hat{H}_{q|i}$.

Marginal Effects of the Latent Class Model

As with all discrete choice models, the marginal effects of a LC models are important behavioral outputs necessary for the interpretation of the parameter estimates and their impact on outcome probabilities. Marginal effects are defined as derivatives of the probabilities that have substantial behavioral

meaning. The posterior estimator of the relevant elasticity is shown in Hensher and Jones (2008):

$$\hat{\sigma}_{km,tj|i} = \sum_{q=1}^Q \hat{H}_{q|i} \hat{\sigma}_{km,j|i|q} \quad (7)$$

An estimator of the average of this quantity over data configurations and firms would be

$$\hat{\bar{\sigma}}_{km,j} = \frac{1}{N} \sum_{i=1}^N \frac{1}{T_i} \sum_{t=1}^{T_i} \hat{\sigma}_{km,tj|i} \quad (8)$$

Variable Selection

A number of explanatory and control variables are introduced from previous research (see Jones and Hensher, 2008; Jones, 2011). Specific financial, market and macro-economic variables tested in the LCM bankruptcy model include:

Leverage. The leverage ratio is measured as total debt to total equity. This variable is central to the debt contracting and accounting choice literatures. Given prior bankruptcy literature, we expect a positive relationship between leverage levels and firm failure.

Operating cash flow to total tangible assets. Operating cash flow to total tangible assets is measured by dividing net operating cash flows (under the direct method) by total tangible assets. Prior to 1992 direct cash flow measures were not available in Australia. Hence, operating cash flow is estimated according to a formula proposed in Lee et al., (1999, p.765).²

EBIT to total assets. EBIT to total assets is measured by earnings before interest and taxation by total assets.

Working capital to total assets. Working capital to total assets is measured by working capital (current assets – current liabilities) by total assets.

²Operating cash flow is computed as $OCE_t = N_t + DAE_t + E_t + G_t + T_t + (CL_t - CL_{t-1}) - (CA_t - CA_{t-1})$ where OCE_t is operating cash flow in year t , NI is earnings before extraordinary items, DAE is depreciation and amortization expense, E is equity in earnings, G is gain (or loss) from sale of long-term assets, T is deferred taxes, CL is current liabilities (less short-term debt), and CA is current assets (less cash and equivalents).

Retained earnings to total assets. Retained earnings is measured by retained earnings divided by total assets.

Quality of earnings. Earnings quality is measured as earnings minus operating cash flow scaled by total assets. The predictive value of this variable in firm failure and fraud detection has been established in Lee et al., (1996) and Rosner (2003).

Cash flow cover. Cash flow cover is measured by dividing operating cash flow by annual interest payments. Higher debt to equity and along with poor capacity to service the debt heightens the financial distress of a firm. A negative relationship between cash flow cover and firm failure is expected.

Total intangible assets to total assets, which has been shown to have an increasing relationship with corporate failure (see Jones, 2011).

New economy sector dummy. A control variable, where a firm belonging to the new economy sector it is coded '1' and zero otherwise. A new economy firm is any firm belong to one or more of the following industry sectors: (i) high technology, (ii) telecommunications, (iii) healthcare and biotechnology, and (iv) internet firms.

Technology crash dummy. A control variable designed to capture the impact of the technology sector collapse in 2001. This variable is coded 1 if the year = 2001 and zero otherwise. This is another control variable which reflects the large number of failures in this sector following the technology crash of 2001.

Recession dummy. A control variable designed to capture the impact of economic recessions over the sampling period. Previous research has identified the importance of macroeconomic factors such as recession on corporate failure. For example, Kane, Richardson and Graybeal (1996) find evidence that recession-induced stress intensifies bankruptcy, for instance by further reducing operating cash flows, thereby impinging on a company's ability to service existing debt. The additive occurrence of stress together with recession-induced stress may push firms across the threshold of insolvency necessary to trigger corporate failure. Recessions are also often characterized by a general loss of

business and consumer confidence, accompanied by growing risk-aversion on the part of lenders and suppliers which may put further pressure on distressed firms leading to a higher incidence of bankruptcy. A technical recession is defined by two successive quarters of negative economic growth. According to the Australian Bureau of Statistics data, over the sample period Australia was formally in recession in 1990, 1991, 1992, 2000 and 2001.³ The year 2001 was not only a recessionary year but the year of a major stock market crash in the technology sector and is picked up by the technology dummy variable. Hence, we tested the impacts of the 1990, 1991, 1992 and 2000 recessions coded as a dummy variable where '1' signifies a recession year and zero otherwise (see Jones, 2011).

Size and Age of the Firm. These are also control variables. Size of the firm is measured as the natural log of total assets. The age variable is represented by a series dummy variables coded '1' if a firm is established 1 year ago or less, and zero otherwise. Five dummy variables were created where the firm is 1, 2, 3, 4, 5 years from establishment. The variable that appears to have the strongest statistical impact on failure was whether a firm was four years or less. This is a dummy variable coded '1' if the firm is less than or equal to four years, and zero otherwise.

Excess value. Excess value equals (market value equity - [book value equity - intangible assets]) divided by market value equity from firm *i* annual report for year *t* balance date (was proposed in Thomadakis, 1977; and later applied in Connolly et al. 1986, see also Jones, 2011). Excess value, being based on market value, was developed by Thomadakis, 1977 as a 'forward looking index of profitability' (p.179). Current market structure is a determinant of future abnormal returns on current and planned investment. It therefore contributes positively to the current market value of the firm (p.180).

³Detailed GDP statistics are provided by the Australian Bureau of Statistics on its website. <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1383.0.55.001Main+Features92009>

Sample Selection

This study models firm failure in two states based on Jones (2011):

State 0: nonfailed firms;

State 1: Failed firms. Failed firms are defined as firms who filed for bankruptcy followed by the appointment of receiver managers/liquidators. This sample includes three major forms of bankruptcy proceeding available under the legislative provisions of the Australian *Corporations Act* (2001): (i) voluntary administration (first introduced in Australia in June 1993 under the *Corporate Law Reform Act*, 1992); (ii) liquidation and (iii) receivership.

Following Jones (2011), we develop two samples for the purposes of model estimation and forecasting. A sample of nonfailed and failed firms was collected between the years 1989 and 2004. Firms were observed to fail or have solvency difficulties at different times over this period. Consistent with previous research, we collect up to five annual reporting periods of data on all firms in categories 0 and 1. The sampling methodology produced a final useable sample for the estimation sample of 2852 firm years, with 1871 firm years in the nonfailed state 0; and 187 firm years in the firm failure sample. The sample of nonfailed firms is drawn over the same time period range as failed firms, and consistent with Jones and Hensher (2004), the proportion of failed to nonfailed firms sampled is approximately equal across each of the years the data are collected.

Data Sources

The firm failure sample was extracted from Huntley's Delisted Company Database, Thomson Reuter's Inactive Database which includes all firms delisted from the ASX between 1989 to 2005. Economic data was obtained from the Australian Bureau of Statistics. Only firms which met the legal definition of failure and had two years of daily stock price data prior to announcement of failure were included in the sample. The bankruptcy announcement date was then

identified using Signal G, a company announcements database maintained by the Australian Stock Exchange pursuant to continuous disclosure requirements.

Empirical Analysis

We estimated a LCM using a number of explanatory and control variables outlined above. As stated above, an important issue in estimating an LCM is specifying number of classes. A 1-class model makes the standard homogeneity assumption that a binary logit model holds true for all cases (the explanatory variables are independent or what is equivalent the IID condition for the error structure). It is crucial to determine the right number of classes – typically, more classes will result in models that better fit the data, but can cause the model to become unstable; but specifying too few could result could ignore important class differences. Typically, a number of models will be estimated on different class number assumptions, and the model fit statistics and significant of the latent class parameters evaluated using different number of classes. We found that the log likelihood function and BIC values improved most when a 2-class model was specified. This model also generated a number of significant latent class parameters. .

Table 1 displays parameter estimates and significance levels across latent classes for a 2-class LCM model. Panel B of Table 1 provides overall model-fit statistics while Panel C displays the within sample classification statistics. We found that a while a 3 or 4 class model improved model fit, individual parameter estimates were less significant overall. The final model was selected based on its overall explanatory and statistical coherence. The 2-class LCM displayed in Table 1 has delivered a very good overall goodness of fit with an adjusted pseudo R^2 of 0.79 (see Panel B). Importantly, when we estimated the model from 1 to 4 classes, there was a significant improvement in the log-likelihood ratio at convergence moving from 1 to 2 classes.

Table 1: Parameter Estimates, t-values, Model Fit Statistics and Classification Performance for the Estimated Latent Class Logit Model

	<i>Class1 Parameter Estimates (t-values in parenthesis)</i>	<i>Class 2 Parameter Estimates (t-values in parenthesis)</i>
Explanatory variables:		
Constant	-0.120(-3.44)	-0.94 (-2.33)
Intangible assets to total assets	.0261(2.67)	.0513(6.41)
Leverage	.061(3.44)	.011(1.99)
EBIT to total assets	-.0712(-2.81)	-.082(-4.15)
Total accruals	.0121(-1.33)	.0034(-1.21)
Net operating cash flow to total assets	-.021(-3.25)	-.0661(-6.13)
Working capital to total assets	-0.046 (-2.43)	-0.081 (-4.63)
Retained earnings to total assets	-.041(-2.56)	-.0861(-4.49)
Cash flow cover	-.02(-2.88)	-.0522(-4.01)
Log of Total Assets	.0252(2.33)	-.092 (-8.12)
Excess value	-.089(-2.44)	-.0822(-6.17)
New economy sector (1,0)	0.016 (3.65)	0.091 (8.13)
Age of firm <=4 years (1,0)	.0561(3.44)	-.088 (5.79)
Technology crash dummy (1,0)	.04 (2.44)	.078 (5.14)
Recession dummy	.001 (.96)	.03 (3.11)
Interaction of Leverage and Voluntary Intangible Asset Capitalization	.0212 (3.12)	.0821 (6.22)
Interaction of Total Accruals and Voluntary Intangible Asset Capitalization	.0351 (3.12)	.0211 (2.87)
Panel B:		
<i>Log-likelihood Statistics</i>		
Log-likelihood at zero	-1165.98	
Log-likelihood at convergence	-112.29	
BIC	2528.15	
X ²	1909.42	
AIC	1467.23	
AIC3	1481.56	
CAIC	1623.34	
Pseudo R square	.79	
Panel C:		
<i>Classification Statistics (Estimation Sample)</i>		
Classification errors	.14	
Reduction of errors	.82	
Entropy square	.79	
Standard R square	.81	
Classification log-likelihood	-3134.06	
AWE	1849.39	

Table 2: Marginal Effects for Latent Class Logit Model

<i>Latent Class Logit Model</i>	
Variables	Marginal Effect on Probability of Firm Failure (t-values)
Voluntary intangible assets to total assets	.051 (4.056)
Leverage	.021 (2.60)
EBIT to total assets	-.027 (-9.3)
Total accruals	.032 (.96)
Net operating cash flow to total assets	-.061 (-9.88)
Working capital to total assets	-.044 (-3.89)
Retained earnings to total assets	-.005 (-1.77)
Cash flow cover	-.016 (-5.69)
Log of Total Assets	-.11 (-8.05)
Excess value	-.0427 (-3.12)
New economy sector (1,0)	.0712 (.874)
Age of firm <=4 years (1,0)	.81 (1.89)
Technology crash dummy (1,0)	.0044 (1.56)
Recession year dummy (1,0) (1992)	.20 (2.21)
Interaction of Leverage and Intangible Assets	.081 (9.88)
Interaction of Total Accruals and Intangible Assets	.092(4.33)

Panel B of Table 1 also reports the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC), Akaike Information Criterion 3 (AIC3), and the Consistent Akaike Information Criterion (CAIC) based on the L^2 and degrees of freedom (df), which the number of parameters in the model. The BIC, AIC and CAIC scores weight the fit and parsimony of the model by adjusting the log likelihood to take into account the number of parameters in the model. These information criteria weight the fit and the parsimony of a model: generally the lower BIC, AIC, AIC3, or CAIC values, the better the fit of the model. We found that BIC score in particular was the most improved for a 2 class model. Further, the R square value improved from .64 for a 1 class model to .79 for a 2 class model.

Within sample classification statistics are also useful for interpreting model performance. When classification of cases is based on modal assignment (to the class having the highest membership probability), the proportion of cases that are expected to be misclassified is reported by the classification. Generally, the

closer this value is to 0 the better; and the model has a relatively low classification error rate of 14%. Reduction of errors (lambda), Entropy R-squared and Standard R-squared are statistics which indicate how well the model predicts class memberships. These statistics are reported on the estimation sample. The closer these values are to 1 the better the predictions as indicated in Table 1 Panel C. Furthermore, AWE is a similar measure to BIC, but also takes classification performance into account. Finally, the classification table cross-tabulates modal and probabilistic class assignments.

Panel A of Tables 1 provides the parameter estimates for each predictor variable in the model; including the degree to which latent classes are statistically different from each other – this is shown by the Wald statistic (which is a test of the null hypothesis that parameter estimates are equal across latent classes). Table 2 provides the marginal effects for the parameter estimated reported in Table 1.

Table 2 provides marginal effects for all explanatory variables reported in Table 1 for the failure outcome. The t -values for marginal effects are in parenthesis. A positive (negative) marginal effect is increasing (decreasing) of the failure outcome. Marginal effects show the increases (decreases) on outcome probabilities when explanatory variables are varied by one unit. The marginal effect for a particular explanatory variable is calculated by taking the derivative (slope) of the probability function while holding all other explanatory variables constant (at their means).

While it is necessary to interpret parameter estimates jointly with the marginal effects, the results in Table 1 indicate many of the explanatory variables appear to have logical signs and are statistically significant. For instance, we would expect a negative coefficient for working capital to total assets, as lower working capital levels are likely to be associated with increased risk of firm failure ($t = -2.43$ and -4.63 for class 1 and 2 respectively). The same observation would be expected for earnings, cash flows and retained earnings to total assets. All variables have negative parameter estimates, indicating that lower levels of these variables are more highly associated with firm failure. Leverage has a positive parameter estimate, indicating that higher leverage increases the risk of firm failure ($t = 3.44$ and 1.99 for class 1 and 2 respectively). Lower levels of excess value also associated with failure ($t = -2.44$ and -6.17 for class 1 and 2 respectively). The interaction of leverage and intangibles is also significant suggesting that higher leverage increases the affects of intangible asset capitalization on corporate failure ($t = 3.12$ and 6.22 for class 1 and 2 respectively). Not all the explanatory variables are significant. For instance, the quality of earnings variable is not significant on either of the latent classes and the recession dummy variable is only significant on latent class 2.

The LCM model in Table 1 has generated a number of statistically significant results. However, there are systematic differences across the two latent classes that are worth mentioning. For instance, latent class 2 (which for convenience we call the ‘small firm’ latent class) shows that firm distress is much more strongly associated with a smaller firms. The t -value of -8.12 is highly significant and negative, suggesting that for class 2,

smaller firms have a much stronger statistical impact on the likelihood of corporate failure. For class 1, which for convenience we describe as the ‘large firm’ latent class, the t -value for the log of total assets is positive (but not significant). Relative to the large firm latent class, the ‘small firm’ latent class also indicates that the new economy dummy, the age of the firm dummy and the technology dummy all have much higher parameter estimates and t -values relative to the large firm latent class (the t -values are 8.13 , 5.79 and 5.14 respectively). The recession dummy also has a stronger statistical impact on the ‘small firm’ latent class ($t = 3.11$), indicating that the recession variable increases the impact of failure for the small firm class. For the large firm class, the parameter is negative but not significant. The small class latent class shows a weaker positive parameter estimate for leverage ($t = 1.99$ vs 3.44 for the large firm class), indicating that while leverage increases the likelihood of failure, the effect of leverage is more pronounced for the large firm class.

Relative to the large firm class, the small firm class also displays a stronger negative parameter estimate for excess value ($t = -6.17$), indicating that negative excess value has a more pronounced impact on the likelihood of failure for the small firm class. The small firm class also shows a stronger negative parameter estimate for working capital to total assets ($t = -4.43$), a stronger negative parameter estimate retained earnings to total assets ($t = -4.49$), a stronger negative parameter estimate for operating cash flows to total assets ($t = -6.13$), and a stronger positive parameter estimate for intangible asset capitalization ($t = 6.41$). This suggests that smaller firms with weaker financial performance and higher rates of intangible asset capitalization increases the likelihood of firm failure. The Wald statistics reported in Table 1 are highly significant, showing that the two latent classes are well separated in terms of their statistical impacts.

Analysis of Marginal Effects

A direct interpretation of the behavioural meaning of parameter estimates reported in Table 1 provides no indication of what the impact of the parameter estimates are on outcome *probabilities*. Furthermore, caution is always needed in interpreting the sign of LCM parameter estimates. Logit parameter estimates

are linear with respect to the utility function, and are *nonlinear* with respect to the outcome probabilities. Nor does the statistical significance of the utility parameter imply the marginal effects are statistically significant. As noted by Jones and Hensher (2004) is possible that the marginal effects of the model have a different sign (and significance) from the parameter estimate. We therefore provide the marginal effects defined as the influence that a unit change in an explanatory variable (or its functional presence) has on the percentage change in the probability of selecting a particular outcome, *ceteris paribus*.

Marginal effects for all significant parameter estimates are reported in Table 2. All parameter estimates appear to have logical and consistent signs. For example, the leverage variable has a positive and significant marginal effect on corporate failure, indicating that a unit increase in this variable *increases* the probability of failure and *decreases* the probability of nonfailure, *ceteris paribus*. In this case, a one unit increase in leverage increases the probability of failure by .021%. A one unit decrease in the working capital to total assets ratio increases the probability of failure by .044%. The technology crash dummy creases the probability of failure by .0044% while membership of the new economy sector increases the probability of firm failure by .0712%. A recession year increases the probability of failure by .20%.

If a firm is 4 years or less, this increases the probability of firm failure by .81%. A one unit decrease in the EBIT to total assets ratio increases the of failure by .027%. A one unit decrease in the retained earnings to total assets increases the of failure by .005%. A one unit decrease in the operating cash flow to total assets increases the probability of failure by .061%. A one unit decrease in the excess value parameter increases the probability of failure by .0427%.

We acknowledge that the economic impact of some of the marginal effects are not that strong. For instance, the excess value variable is still quite small in absolute terms. However, it needs to be borne in mind that our sample is based on failure frequency rates that are much closer to actual failure rates observable in practice. Our model's marginal effects tend to be smaller in absolute terms because they are derived from probabilities and parameters

estimates which are based on a *very* high proportion of non-failures relative to firm failure (i.e., a much larger change in a marginal effect is needed to move a company (in probability terms) from the non-failure category to the firm failure category). Furthermore, as many failed and distress firms in our sample tend to have very small market capitalizations (as well as very thin trading liquidity), large changes in the market value of equity will not necessarily have a significant impact on financial distress levels.

Out of Sample Forecasting Accuracy of the LCM Model

Having evaluated the model-fit information, parameter estimates and marginal effects, we now turn to the prediction outcomes. Calculating probability outcomes for a LCM is considerably more simple than for open form solution models (such as mixed logit) as it is a closed form model. In deriving the probability outcomes, we calculated the probabilities from equations above on the holdout sample. Consistent with the approach adopted in the discrete choice literature, we focus on a sample enumeration method which recognizes that the estimated model is based on a sample drawn from a population and the application of the model must preserve the full distribution of information obtained from the model system (see Train, 2003). This includes the outcome probabilities. Thus we aggregate the probabilities associated with each outcome across the entire sample to obtain the predicted values. Implementing a sample enumeration strategy on our hold out sample, we can evaluate the predictive performance of the latent class model.

The latent class logit model has a high level of predictive accuracy on a holdout sample across the nonfailure and failure alternatives. The latent class model is 86% accurate in predicting the nonfailure outcome (84.5% actual vs. 85.2% predicted), and 77% accurate in predicting the failure outcome. The model is accurate up to three years prior to failure. One year prior to failure, the model is 91% accurate in predicting the nonfailure outcome 84% accurate in predicting the failure outcome. Two years prior to failure, the model is 85% accurate in predicting the nonfailure outcome 79% accurate in predicting the failure outcome. Three years prior to failure, the

model is 77% accurate in predicting the nonfailure outcome 74% accurate in predicting the failure outcome.

Notwithstanding the relatively strong predictive accuracy of the LCM, it needs to be acknowledged that selecting a model based solely on prediction capability of a hold out sample is to deny the real value of models in evaluating the behavioural responses in the market to specific actions, planned or otherwise, as represented by the elasticities linked to specific explanatory variables.

Elasticities are arguably the most important behavioural outputs, although confidence in sample-based predictions of state shares adds to the overall appeal of an empirical model as a policy tool. A behaviourally relevant model should be able to predict with confidence what is likely to happen when one of more explanatory variables take on new values in real markets.

Conclusions

The LCM framework is adopted in this study because of its capacity to handle highly restrictive statistical assumptions which can distort parameter estimates and probability outcomes, and consequently lead to misleading interpretations of model outputs. LCMs also provide a significant amount of additional behavioral information useful for evaluating the role and influence of explanatory variables on outcome probabilities. The estimated LCM for this study is assessed on a number of stringent criteria including: (i) the strength of the model fit statistics, (ii) the sign and significance of the parameter estimates and marginal effects and (iii) predictive accuracy on a holdout sample. The LCM performs significantly better than a standard logit model (which is equivalent to a single class LCM) both in terms of model fit statistics, classification performance within sample and holdout predictions. LCMs are particularly valuable to current research and practice as they share many of the benefits of a mixed logit model but are significantly easier to estimate and interpret in an applied bankruptcy setting.

References

- Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, 23(4): 589-609.
- Altman, E., Haldeman, R. and Narayan, V. (1977). ZETA Analysis: a new model to identify bankruptcy risk of corporations, *Journal of Banking and Finance*, June, 1(1):29-54.
- Altman, E. (2001). *Bankruptcy, Credit Risk and High Yield Junk Bonds*, Blackwell Publishers, New York, NY, USA.
- Australian Accounting Standards Board. (1992). *Approved Australian Accounting Standard AASB 1026: Statement of Cash Flows*. Australian Accounting Research Foundation, Melbourne, Victoria, Australia.
- Australian Corporations Act (2001). Cwt, Canberra, ACT, Australia. (<http://www.asic.gov.au/asic/ASIC>),
- Australian Stock Exchange (2003). *Listing Rules*, Sydney, NSW, Australia. <http://www.asx.com.au/about/l3/ListingRules>
- Australian Stock Exchange (2003). *Market Comparative Analysis*, Sydney, NSW, Australia. (<http://www.asxtra.asx.com.au>).
- Bahnsen, P., and Bartley, J. (1992). The sensitivity of failure prediction models, *Advances in Accounting*, 10(1): 255-278.
- Ben-Akiva, M., Bolduc, D. and Walker, J. (2001). *Specification, identification and estimation of the logit kernel (or continuous mixed logit) model*, MIT Working paper, Department of Civil Engineering, Massachusetts Institute of Technology, Boston, MASS, USA.
- Bharath, S. T. and Shumway, T. (2004). *Forecasting default with the KMV-Merton model*, AFA 2006 Boston Meetings Paper., Boston, MASS, USA. Available at SSRN: <http://ssrn.com/abstract=637342>
- Bhat, C.R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed

- multinomial logit model, *Transportation Research Part B*, 35(1), 677-693.
- Brownstone, D. and Train, K. (1999). Forecasting new product penetration with flexible substitution patterns, *Journal of Econometrics*, 89(2), 109-129
- Bulow, J. I. and Shoven, J. B. (1978). The bankruptcy decision, *Bell Journal of Economic*, 9 (2): 437-456.
- Casey, C. and Bartczak, N. (1985). Using operating cash flow data to predict financial distress: Some extensions, *Journal of Accounting Research*, 23(1): 384-401.
- Clark, K., and Ofek, E. (1994). Mergers as a means of restructuring distressed firms: and empirical investigation, *Journal of Financial and Quantitative Analysis*, 29(4): 541-565.
- Clark, T. A. and Weinstein, M. I. (1983). The behaviour of the common stock of bankrupt firms, *Journal of Finance*, 38(2): 489-504.
- De Silva, R., Nguyen, T. and Walter, T. (2004). Market returns to acquirers of substantial assets, *Australian Journal of Management*, 29(1): 111-134.
- Delaney, K. (1991). *Strategic Bankruptcy: How Corporations and Creditors Use Chapter 11 to their Advantage*. University of California Press, Berkeley, California, USA.
- Dichev, I. D. (1998). Is the risk of bankruptcy a systematic risk?, *Journal of Finance*, 53(3): 1131-1148.
- Duffie, D., and Singleton, K. (2003). *Credit Risk: Pricing, Measurement and Management*. Princeton University Press, Princeton, NJ, USA.
- Duffie, D., Saita, L. and Wang, K. (2005). *Multi-period Corporate Default Prediction with Stochastic Covariates*. Working Paper, Graduate School of Business, Stanford University, San Francisco, CAL, USA.
- Frino, A., Jones, S., and Wong, J. (2005). *Market Behaviour Around Bankruptcy Announcements: Evidence From The Australian Stock Exchange*. Working Paper, Discipline of Finance, The University of Sydney, Sydney, NSW, Australia.
- Gentry, J., Newbold, P. and Whitford, D. (1985). Classifying bankrupt firms with funds flow components, *Journal of Accounting Research*, 23(2): 146-160.
- Goodman, L. (2002). Latent class analysis: The empirical study of latent types, latent variables, and latent structures, *Applied Latent Class Analysis*, edited by J. A. Hagenaars and A. L. McCutcheon, Cambridge University Press, Cambridge, England.
- Gopinath, D., Schofield, M., Walker, J. and Ben-Akiva, M. (2004). Comparative analysis of discrete choice models with flexible substitution patterns, paper prepared for the 2005 Annual meeting of the Transportation Research Board, Washington D.C, USA.
- Gourieroux, C., and Monfort, A. (1996). *Simulation-Based Methods Econometric Methods*. Oxford University Press, Oxford, England.
- Greene, W. H. and Hensher, D.A. (in press). Mixed logit and heteroscedastic control for random coefficients and error components, *Transportation Research E. Transportation Research Part E: Logistics and Transportation Review*, 43 (5): 610-623.
- Greene, W. H., Hensher, D.A. and Rose, J. (2006). Accounting for heterogeneity in the variance of unobserved effects in mixed logit models, *Transportation Research B: Methodological - An International Journal*, 40(1): 75-92.
- Hensher, D.A. and Greene, W.H. (2003). Mixed logit models: State of practice, *Transportation*, 30 (2). 133-176.
- Hensher, D.A., Rose, J. and Greene, W. H. (2005). *Applied Choice Analysis: A Primer*, Cambridge University Press, Cambridge.
- Hill, N., S. Perry, and S. Andes. (1996). Evaluating firms in financial distress: an event history analysis. *Journal of Applied Business Research*, 12(1):60-71.
- Hribar, P. and D. Collins. (2002). Errors in estimating accruals: implications for empirical research, *Journal of Accounting Research*, 40(1):105-134.

- Jones, S. and Hensher, D.A. (2004). Predicting firm financial distress: A mixed logit model, *The Accounting Review*, 79(4): 1011-1038.
- Jones, S. and Hensher, D.A. (2007). Modelling corporate failure: A multinomial nested logit analysis for unordered outcomes, *British Accounting Review*, 39(1): 89-107
- Jones, S. and Hensher, D.A. (2008). *Advances in Credit Risk Modelling and Corporate Bankruptcy Prediction* (Editors), Cambridge University Press, Cambridge, United Kingdom.
- Jones, S. (2011). Does the capitalization of intangible assets increase the predictability of corporate failure, *Accounting Horizons*, 25(1): 41-70
- Kane, G.D., Richardson, F.M., Graybeal, P. (1996). Recession-induced stress and the prediction of corporate failure, *Contemporary Accounting Research*, 13(2): 631-650.
- Kealhofer, S. (2003). Quantifying credit risk 1: Default prediction, *Financial Analysts Journal*. 59(1): 30-44.
- Lau, A.H. (1987). A five-state financial distress prediction model, *Journal of Accounting Research*, 25(1): 127-138.
- Lazarsfeld, P. F. and Henry, N. W. (1968). *Latent Structure Analysis*, Houghton Mifflin, Boston, MASS, USA.
- Lee, T.A., Ingram, R.W. and Howard, T.P. (1996). The difference between earnings and operating cash flow as an indicator of financial reporting fraud, *Contemporary Accounting Research*, 16(4): 749-86
- Louviere, J.J., Hensher, D.A. and Swait, J.F. (2000). *Stated Choice Methods and Analysis*, Cambridge University Press, Cambridge, England.
- McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response, *Journal of Applied Econometrics*, 15(2): 447-470.
- Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, 18(1): 109-131.
- Opler, T., and Titman, S. (1995). *Financial Distress and Capital Structure Choice*, Working Paper, Boston College Boston, MASS, USA.
- Pastena, V. and Ruland, W. (1986). The merger/bankruptcy alternative, *The Accounting Review*, 61(2):288-301.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: a simple hazard model, *Journal of Business*, 74(1): 101-124.
- Stern, S. (1997). Simulation-based estimation, *Journal of Economic Literature*, 35: 2006-2039.
- Sutton, R. I. and Callaghan, A. L. (1987). The stigma of bankruptcy: spoiled organizational image and its management. *Academy of Management Journal*, 30(1): 405-436.
- Train, K. (2003). *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, England.
- Vassalou, M., and X. Yuhang (2004). Default risk in equity returns, *The Journal of Finance*, April, 59(2):831-868.
- Ward, T. (1994). An empirical study on the incremental predictive ability of Beaver's naïve operative flow measure using four-state ordinal models of financial distress, *Journal of Business Finance & Accounting*, 21(2): 547-562.
- Zmijewski, M. (1984). Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research*, 22: Supplement: 59-82

