WORKING PAPER No. 6
January 1992

SMALL FIRM BANKRUPTCY PREDICTION:
SPAN AND LOCATION OF TREND EFFECTS

Robert Cressy

Warwick Business School's Small and Medium Sized Enterprise Centre Working Papers are produced in order to make available to a wider public, research results obtained by its research staff. The Director of the CSME, Professor David Storey, is the editor of the series. Any enquiries about this series, or concerning the research undertaken within the Centre, should be addressed to:

The Director, CSME
Warwick Business School
University of Warwick
Coventry CV4 7AL
Tel [01203] 523741/523692
Fax [01203] 523747
SMALL FIRM BANKRUPTCY PREDICTION: 
SPAN AND LOCATION OF TREND EFFECTS

by
Robert Cressy, University of Warwick, England

1 INTRODUCTION
Much work has been done in the last two decades to
estimate the determinants of bankruptcy of large firms. 
(See Storey, Keasey, Watson and Wynarzcyck (henceforth 
SKWW) 1987 for references). Very little apart from the 
early work of Edminster (1972) in the US and more 
recently the work of SKWW (1987) in the UK appears to 
have been attempted in the area of small firm bankruptcy. 
In a companion paper (Cressy 1992) I examined the 
influence of financial trends, macro- and industry-
effects on small firm bankruptcy potential and showed 
that a financial ratio trends model performs as well as 
any model incorporating these other effects. In the 
present paper I analyse the fit of a number of financial 
trend models of varying 'location' (temporal relation of 
data to event date) and 'span' (number of lags in data). 
Three-, four- and five-year lag structures are modelled 
and in each case the fit is examined for periods selected 
at different calendar dates prior to the event date.

Our results confirm the conclusion of the companion 
paper that several years' date on financial ratio 
variables are required to provide reasonable predictive 
accuracy on small firm solvency rather than the 
one year's accounts information traditionally used 
in large firm analysis. However in the current paper 
we are able to quantify the degree to which goodness-of-
fit is affected by both the span of the dataset used 
and by its location. Some rather interesting results 
are obtained. In particular we find that the best 
model explains about 60% of the variation in
bankruptcy probabilities across firms and that the accuracy of the models does not decline monotonically as the distance from the event date increases.

II THE RELEVANT LITERATURE

Sample Selection Issues
The seminal work of Altman (1968) spawned a huge literature on the prediction of bankruptcy. (See e.g. Zavygren (1983) for a survey). However much of this work has recently been placed into question by two perceptive papers (Zmijewski(1984) and Palepu(1986)). We have summarised the basic results of these papers and their implications for small firm bankruptcy analysis in particular in another paper (Cressy, 1991) and will not repeat them here. Suffice to say that samples used to estimate bankruptcy functions must either be chosen at random from the population or if nonrandomly selected estimated parameters must be adjusted to take known parameter biases into account.

Choice of Cutoff Probability
A cutoff value in the logistic model serves to divide the population into two groups on the basis of probabilities. If a firm's estimated probability is above the cutoff it is classified as bankrupt and if below as active. Intuitively the choice of cutoff point must reflect the expected gains/losses over alternatives. Unfortunately little attention has been paid in the literature to the optimality criteria for such cutoffs.¹

One method of choosing cutoffs is to minimise the Savage(1954) regret or loss function: Suppose we can evaluate the economic losses to the decision-maker arising from Type I and Type II errors
in classification. Imagine starting with an arbitrary probability cutoff value $p_0$. Firms are classified as bankrupt or active as their observed probability $p_i$ falls above or below $p_0$. This rule generates an expected loss with weights the posterior joint probabilities of pairs $(p_i, y_i)$, $y_i=0$ or $1$, where $y_i$ is the bankruptcy variable such that a value of one indicates a bankrupt firm and a value of zero an active firm. If the cutoff value is varied marginally the Type I/II error will in general rise (or fall) and the Type II/I error fall (or rise), changing the expected loss accordingly. If the expected loss falls it follows that we have not chosen $p_0$ optimally.² Palepu (1986) has demonstrated the classification biases that result from such omission. In general bankrupt classification accuracy tends to be overstated and active classification accuracy understated.

It is thus important to choose the cutoff probability in logistic regression according to a suitable criterion. If possible economic theory should be of direct assistance as in Palepu (1986) where the choice of cutoff follows from an economic model of the takeover process. However since Palepu's theory presupposes a competitive bidding process for the firm's equity it is not applicable to the analysis of small firms and in particular to the analysis of small firm bankruptcy.³ We therefore adopt the criterion, familiar to statisticians, of expected loss minimisation where losses are assumed to be positive and equal for Type I and Type II errors and zero elsewhere.

More formally, denoting the estimated probability of bankruptcy by $p_i$, the set of firms classified into the bankrupt category by $\mathcal{P}^* = \{p_i: p_i \geq p^*\}$ and the loss from action $a_i$ when state $s_j$ holds by $L(a_i, s_j)$ the expected loss from the classification scheme is given in general by
EL = \sum_{i} L(p^*_i, i \in A) + \sum_{i} L(p^*_i, i \in B)

= \sum_{i} \int_{p^*_i} f(p^*_i|y=0) dp^*_i q_0 + \int_{p^*_i} f(p^*_i|y=1) dp^*_i q_1

(1)

where \( L_{BA}, L_{AB} \) is abbreviated notation for misclassification losses and \( q_0, q_1 \) are prior probabilities of being an A or a B respectively. Minimising this expected loss or regret with respect to \( p^*_i \) we get

\[
\frac{f(p^*_i|y=0)}{f(p^*_i|y=1)} = \frac{L_{AB} q_1}{L_{BA} q_0}
\]

subject to

\[
\frac{f'(p^*_i|y=0)}{f'(p^*_i|y=1)} \leq \frac{L_{AB} q_1}{L_{BA} q_0}
\]

Thus at the optimum the marginal expected losses from misclassification based on posterior densities must be equal.\(^4\)

Equation 2 specialises of course to

\[
\frac{f(p^*_i|y=0)}{f(p^*_i|y=1)} = \frac{q_1}{q_0}
\]

when we assume losses are symmetric and equal to one. The problem is in fact identical to that of minimising prediction error or maximising prediction accuracy.\(^5\)

We note that the interpretation of equation 4 is very simple for single-peaked distributions and amounts to locating the point of intersection of one posterior distribution and a scaled up/down version of the other. Multi-peaked distributions require further analysis and are best examined graphically. However space does not
Choice of Prior Probabilities
Another important choice in the determination of $p^*$ is that of the prior probabilities $q_j$. In the present work we follow the standard procedure in the statistical literature of identifying prior probabilities with population proportions of B's and A's respectively.

The Small-Firm Literature
Very little work has been done in the area of predicting small firm bankruptcy. Edmister (1972) on US data and SKWW (1987) on UK data are apparently the only published econometric studies currently available. We provide here a brief summary of their results.

Edmister (1972)
Edmister used discriminant analysis in the vein of Altman to examine the bankruptcy characteristics of 42 US small firms. Two analyses were performed, one using firms for whom one year's financial data was available prior to the event date and another using firms for whom three years' financial data was available. Sampling rates were 100% of the bankrupts in each subsample and 15% and 19% of the actives for the one- and three-year sets respectively. He employed a stepwise approach to obviate problems of multicollinearity. Edmister found that intertemporal instability displayed by small firm financial ratios made the model with one year's data useless for bankruptcy prediction into the holdout sample. He concluded that one should use three-year averages instead of annual data and the three-year model so structured predicted well in the holdout sample.

Edmister's study suffers from estimation bias due to the sample
selection procedure employed (Palepu, 1986). The averaging method used in the three-year analyses also has the econometric disadvantage that it imposes unnecessary parameter restrictions on the model being estimated. Thus using a three year average assumes that the regression weights attached to each year's variable value are the same. A general lag structure in which the weights are allowed to be determined by the data is clearly a better specification.

SKWW (1987)

SKWW (1987) in a pioneering study of UK small firm bankruptcy used discriminant and logit analysis to examine the bankruptcy potential of a set of 636 small UK firms. Our discussion of their extensive study will relate primarily to their logit analysis, the technique employed in the present paper.

The SKWW sample was defined as all firms in the Northern region of England with at least one years published (Companies House) accounts data and with less than 200 employees. Firms had in addition to be Limited, single-plant, independent, manufacturing companies. The 8 variables used in logit analysis were selected from a larger set of 12 variables by factor analysis to represent the characteristics of liquidity, profitability, and so on thought to be relevant to the prediction of bankruptcy. The set of 12 variables was as follows:

W1. Current assets/ Current liabilities
W2. Net profit/Total assets
W3. Fixed assets/Total assets
W4. (Pre-tax profit+depreciation)/Total debt
W5. (Pre-tax profit before directors' fees + interest)/Total debt
W6. (Total debt excl. bank overdraft)/Total assets
W7. Current assets/Total assets
W8. (Current assets-stock)/Total assets
W9. Net profit/Fixed assets
W10. Fixed assets/Net worth
W11. Net profit/(Current assets-current liabilities)
W12. Pre-tax profit/Net worth

Eight variables W1, W2, W4, W5, W7, W9, W11, W12 were selected by SKWW from the above set by means of factor analysis and a series of 5 annual logit regressions ([t-k+t], k=1,..5) were run, where t is the event date (insolvency/non-insolvency) and t-k refers to the data year used to predict. Two sets of regressions were run using firms with 3 and 7 consecutive years' 'complete' (8-variable) data respectively.

All the one-year regressions were significant at the 5% level but only one or two individual variables were significant in each regression. Sets of variables that were significant changed over time. Some years would have liquidity, others measures of asset structure and yet others profitability as the significant explanatory variables in the model. Classification accuracy was no better than MDA on the same dataset and for bankrupt firms this varied between 7 and 100% on the original sample. Cutoffs were chosen to minimise the $\chi^2$ of the model. No $R^2$ statistics were reported.

A number of problems arise in interpreting the results of the SKWW study most significant of which is the sample selection procedure followed and the absence of any adjustment factor in the estimation process. These have been discussed in Cressy(1991) and we shall not repeat them here. Suffice it to say that the classification accuracy and $R^2$ falls substantially if the selection
bias is corrected for.

II THE PRESENT STUDY

To avoid some of the problems of bias outlined above the current empirical analysis will be conducted on the population of firms in the Warwick Small Firms database (SKWW, 1987). The data are pooled to increase the number of bankrupt firm observations in the sample. The logic behind this is as follows.

The (average) sample probability of failure in a specific year $t$ is given by

$$p_B(t) = \frac{n_B(t)}{n(t)}$$

where $n_B(t) =$ number of bankrupts in year $t$

$n(t) =$ total number of firms (bankrupts plus actives) in $t$

The (average) sample probability of failure in any year is obtained by summing over $t$:

$$p_B = \frac{\Sigma n_B(t)}{\Sigma n(t)}$$

This procedure thus allows for the fact that it is not possible to distinguish ex ante between Bs and As. Thus an observation on a bankrupt firm say two years before bankruptcy is treated as one on an active firm in that year. Likewise active firms will in general appear more than once in the observation set since they are potential bankrupts in these years. The result is a dataset that has many more observations than the 636 firms in the SKWW study.
Because of the misleading nature of classification accuracy as a measure of fit or of prediction accuracy (Cressy, 1991) we shall rely primarily on the pseudo-\(R^2\) as a measure of how well our models 'explain' the data. (See Maddala, 1983).\textsuperscript{11}

Before moving on to the empirical analysis proper we introduce four new variables into the SKWW set on the basis of their popularity in the financial and accounting literature. They are defined as follows:

W13. Cashflow/Total debt  
W14. Equity/Total debt  
W15. Quick assets/Current liabilities  
W16. Creditors/Debtors

Apart from W14 which measures in some sense the financial risk of the firm's activities these additional variables are indicative of the degree of liquidity of the firm's balance sheet.

III Financial Trend Regressions

The object of the analysis is to evaluate the performance of financial trend models both in terms of the distance of the data from the event date and in terms of the number of lags included in the specification. The sample used in these regressions is the population of firms with complete\textsuperscript{12} observations on current and lagged values of W1-W16 defined above\textsuperscript{13}.

Initial experiments with the use of all variables displayed a very high degree of collinearity in the explanatory variables. To obviate this we used (following Edmister and SKWW) a stepwise procedure. All possible regressions with three-, four- and five-year lag structures for all possible calendar dates prior to
event dates were run. Tables 1-4 provide the results for the three-to six-year lag regressions respectively. (In the Tables the terminology "Regression 20" means the regression using variables L2Wi, LWi, and Wi; "Regression 31" means the regression using L3Wi, L2Wi, LWi; etc etc.)
Table 1: 3-Year Trend Regressions

Regression 20

<table>
<thead>
<tr>
<th>Coeff</th>
<th>-4.5</th>
<th>-.28</th>
<th>-.44</th>
<th>-.65</th>
<th>-.25</th>
<th>.73</th>
<th>.49</th>
<th>-1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig Lvl(%)</td>
<td>.00</td>
<td>.22</td>
<td>.01</td>
<td>.06</td>
<td>1.76</td>
<td>.04</td>
<td>1.02</td>
<td>.33</td>
</tr>
</tbody>
</table>

N=2045  N =2004  N =41

\[ MR^2 = 0.294 \]  (Sig Lvl(%)=.00)  Optimal cutoff = 0.31

% Classification Accuracy  % Prediction Accuracy

<table>
<thead>
<tr>
<th>B</th>
<th>A</th>
<th>T</th>
<th>B</th>
<th>A</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>100</td>
<td>98</td>
<td>100</td>
<td>98</td>
<td>98</td>
</tr>
</tbody>
</table>

Regression 31

<table>
<thead>
<tr>
<th>Coeff</th>
<th>-3.8</th>
<th>-.44</th>
<th>-.13</th>
<th>-.19</th>
<th>-.12</th>
<th>.49</th>
<th>.13</th>
<th>-.38</th>
<th>.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig Lvl(%)</td>
<td>.00</td>
<td>.00</td>
<td>4.7</td>
<td>2.1</td>
<td>2.3</td>
<td>.04</td>
<td>8.0</td>
<td>3.3</td>
<td>4.5</td>
</tr>
</tbody>
</table>

N=2062  N =2004  N =58

\[ MR^2 = 0.231 \]  (Sig Lvl(%)=.00)  Optimal cutoff = 0.53

% Classification Accuracy  % Prediction Accuracy

<table>
<thead>
<tr>
<th>B</th>
<th>A</th>
<th>T</th>
<th>B</th>
<th>A</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>100</td>
<td>97</td>
<td>80</td>
<td>97</td>
<td>97</td>
</tr>
</tbody>
</table>
Regression 42

Coeff  -3.9  -.53  .15  .11  -.57  -.60  -  .15  .59  -.46  -.15  -.55
Sig Lvl(%) .00  .00  1.4  3.4  .00  .01  7.47  .04  .18  3.8  .11

N=2072  N_A=2004  N_B=68

MR^2 = 0.342  (Sig Lvl(%)=.00)  Optimal cutoff = 0.47

% Classification Accuracy   % Prediction Accuracy
  B  A  T  B  A  T
10  100  97  70  97  97

Regression 53
C L4W2  L3W10  L5W10

Coeff  -3.3  -.24  .15  .19
Sig Lvl(%) .00  .21  1.0  2.1

N=2063  N_A=1988  N_B=75

MR^2 = 0.165  (Sig Lvl(%)=.00)  Optimal cutoff = 0.23

% Classification Accuracy   % Prediction Accuracy
  B  A  T  B  A  T
 4  100  96  50  96  96
Table 2: 4-Year Regressions

### Regression 30

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>L2W2</th>
<th>W9</th>
<th>L2W9</th>
<th>L3W10</th>
<th>LW9</th>
<th>L2W7</th>
<th>L2W10</th>
<th>L3W11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td>-4.7</td>
<td>-0.56</td>
<td>-0.58</td>
<td>-0.45</td>
<td>-0.19</td>
<td>-0.35</td>
<td>0.51</td>
<td>0.29</td>
<td>0.67</td>
</tr>
<tr>
<td>Sig Lvl(%)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3.2</td>
<td>5.5</td>
<td>4.3</td>
<td>4.1</td>
<td>3.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

N = 1641  \( N_A = 1611 \)  \( N_B = 30 \)

\[
MR^2 = 0.426 \quad (\text{Sig Lvl(%) =} 0.00) \quad \text{Optimal cutoff} = 0.18
\]

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>33</td>
<td>99</td>
</tr>
</tbody>
</table>

### Regression 41

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td>-4.5</td>
<td>-0.65</td>
<td>0.25</td>
<td>-0.93</td>
<td>-0.93</td>
<td>-0.29</td>
<td>0.39</td>
<td>-0.50</td>
<td>-1.09</td>
<td>0.62</td>
</tr>
<tr>
<td>Sig Lvl(%)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.7</td>
<td>2.1</td>
<td>0.04</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 1641  \( N_A = 1611 \)  \( N_B = 30 \)

\[
MR^2 = 0.447 \quad (\text{Sig Lvl(%) =} 0.00) \quad \text{Optimal cutoff} = 0.31
\]

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>99</td>
</tr>
</tbody>
</table>
Regression 52

Coef -4.2  .66  -.20  -1.31  -.42  .53  -.55  -.35  .25  -.51
Sig Lvl(%) .00  1.1  .46  .00  .07  1.8  .13  1.6  1.6  .11

$N = 1662 \quad N_A = 1611 \quad N_B = 51$

$\text{MR}^2 = 0.389 \quad (\text{Sig Lvl(%) = .00}) \quad \text{Optimal cutoff} = 0.30$

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3: 5-Year Regressions

Regression 40

<table>
<thead>
<tr>
<th>C</th>
<th>L2W2</th>
<th>W9</th>
<th>L4W2</th>
<th>L4W3</th>
<th>L3W2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td>-4.97</td>
<td>-.94</td>
<td>-.64</td>
<td>-.43</td>
<td>-.87</td>
</tr>
<tr>
<td>Sig Lvl(%)</td>
<td>.00</td>
<td>.00</td>
<td>.23</td>
<td>.00</td>
<td>.72</td>
</tr>
</tbody>
</table>

N=1291    N=1268    N=23

MR^2 = 0.495 (Sig Lvl(%)=.00)  Optimal cutoff = 0.18

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>35</td>
<td>100</td>
</tr>
</tbody>
</table>

Regression 51

<table>
<thead>
<tr>
<th>C</th>
<th>L2W9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td>-3.8</td>
</tr>
<tr>
<td>Sig Lvl(%)</td>
<td>.00</td>
</tr>
</tbody>
</table>

N=1299    N=1268    N=31

MR^2 = 0.213 (Sig Lvl(%)=.00)  Optimal cutoff = 0.19

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 4: 6-Year Regression

<table>
<thead>
<tr>
<th>Coeff</th>
<th>6.7</th>
<th>-0.97</th>
<th>-0.98</th>
<th>0.66</th>
<th>-1.0</th>
<th>-0.69</th>
<th>-0.79</th>
<th>-2.5</th>
<th>-0.89</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig Lvl(%)</td>
<td>0.00</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>3.04</td>
<td>9.4</td>
<td>6.18</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ N=993 \quad N=976 \quad N=17 \]

\[ MR^2 = 0.586 \quad (\text{Sig Lvl}(%)=.00) \]

Optimal cutoff = 0.16

<table>
<thead>
<tr>
<th>% Classification Accuracy</th>
<th>% Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>71</td>
<td>99</td>
</tr>
</tbody>
</table>

Notes: (i) Our measures of goodness-of-fit are McFadden's \( R^2 \) adjusted for degrees of freedom defined as \( MR^2 = (\chi^2(p)-2p)/(-2\log(L_\omega)) \) where \( L_\omega \) is the maximum likelihood function under the Null hypothesis and \( p \) the number of parameters estimated in the Alternative hypothesis.
General Discussion: 3-Year Regressions

We note that overall sample sizes are large (greater than 2000) with numbers of bankrupts substantial at between 41 and 75. All of the 3-year regressions are significant at below the 1/100% level. However the $R^2$'s vary from 16% (reg 53) to 34% (reg 42) with an average of 26%. Thus none of the regressions explains a very high proportion of the variation of probabilities across firms. B-classification accuracy varies between 4% and 10% with an average of 6.5%. B-prediction accuracy varies between 50% and 100% with an average of 75%. Thus the variation in $R^2$ is to some extent mirrored in the B-classification and prediction accuracies.

We note that there is a great disparity between the average prediction accuracy and classification accuracy of the models, with generally inverse relationship observable in the sample.

Individual coefficients are despite the poor fit of the models highly significant with the majority significant at well below 1%. Numbers of significant variables range from as few as three (reg 53) to as many as ten (reg 42).

Of the variables that appear in the final equations profitability (LiW2) appears as significant in every equation, often figuring three periods prior to the event date (as L2W2) where it is negatively related to the probability of insolvency. Capital adequacy (LiW10) also figures prominently usually four periods prior to the event date (as L3W10) where it is negatively related to the probability of insolvency. Otherwise there is very little overlap in the variables for the 3-year regressions.

We note that there is no monotonic decrease/increase in the $R^2$ or B-classification accuracy as one moves away from the event date.
This finding clearly differs from the large firm results that almost universally suggest that fit declines as one moves away from the event date.\textsuperscript{16}

Finally the best-fitting model from the 3-year set is reg 42 with ten variables and an $R^2$ of 34%.

General Discussion: 4-Year Regressions

The overall sample size is still very large but falls from over 2000 to around 1650 due to missing data. Numbers of bankrupts are still respectable at between 30 and 51. All three regressions are again significant at below the 1/100% level. $R^2$.s are on average higher than in the three-year regressions at around 42%, with a range of 39% to 45%. B-prediction accuracy is on average higher than for the 3-year regressions at 54% with a range of 50 to 60%. B-classification accuracy is also higher with an average of 23% and a range of 18% to 33%. Thus there is much less variation in $R^2$ between the three regressions than there is in the classification-based measures of goodness-of-fit commonly used.

Of the variables that are significant in the optimal 4-year regressions we note that once again profitability (LiW2,LiW9,LiW10) plays a recurring role usually three or more years prior to the event date.

The best-fitting model of the 4-year regressions is reg 41 with an $R^2$ of 45%. However model 30 has an $R^2$ of 43% and is considerably simpler. The latter has nine variables though when one realises these are often lagged values of the same variables the number effectively reduces to five. The most regularly occurring variables here are LiW9 and LiW10 representing measures of short- and long-run profitability respectively. LiW2, another measure of
profitability, is also highly significant. Thus we recommend this as the operational equation in the subset.

General Discussion: 5-Year Regressions
The sample size drops from about 1600 to about 1300 with numbers of bankrupts now between 23 and 31. Both regressions are significant at below 1/100%. The average $R^2$ is at 35% lower than for the 4-year set but the maximum rises above that of the 4-year regressions to 50%. The average classificatory and prediction accuracies are 20% and 67% respectively.

Regarding the variables of interest in the best 4-year model (reg 40) we note that once more profitability dominates. In fact apart from capital intensity (LiW3) only profitability (LiW2,LiW9) matters in reg 40 and generally operates with the 'expected' sign.

General Discussion: 6-Year Regression
The $R^2$ for the 6-year model model is 58.1% and significant at below the 0.5% level. Five of the seven variables in the model are significant at 1% or below. Individual partials vary between 11% and 38%. Seventy one percent of failures are identified at relatively small cost in terms of Type I errors. The failure prediction accuracy of the model at 50% is satisfactory.

The $R^2$ for reg 50 at 58% is greater than for any other model, including reg 40 with an $R^2$ of 50%. The latter however has the virtue of a smaller number of variables and smaller data requirements. However the difference in explanatory power could be sufficient to justify its choice in certain contexts (where e.g. data availability is not a problem).
Interpretation

Profitability (Net profits/Total assets, LiW2) appears in all models and in virtually every prior year as a statistically significant determinant of bankruptcy. It has moreover a quite high negative correlation with the bankruptcy variable (up to 35% for 3 years' prior) and has obvious intuitive appeal.

Liquidity figures in the early years (L5W5, L5W7, L5W11⁻¹) as a determinant and as expected reduces the probability of bankruptcy. However it is overtaken by profitability (L4W2,L3W12,L2W2,W2) in the later years as the primary determinant of solvency.

No other factors play so dominant role as profitability and liquidity in the small company's financial fortunes. They must therefore be regarded as the major determinants of bankruptcy for small firms and in which the influence of trends is paramount.¹⁷

There are two other ratios worthy of mention. (LiW1,(LiW3)⁻¹)¹⁸ are positively correlated with bankruptcy. Interpreted as measures of liquidity this result may seem counterintuitive, although it is possible with some ingenuity to rationalise the phenomenon. However it perhaps more plausible to interpret these ratios as measures of capital intensity. Under this interpretation the results indicate that more capital-intensive firms are less bankruptcy-prone.

It is also of interest to note that the financial economist's darling, the debt-equity ratio, and the accountant's favourite, the quick ratio, play no role in the reg 50 model. This suggests that they have no independent explanatory power in the determination of solvency probabilities. Thus although a higher debt-equity ratio may increase the likelihood of bankruptcy if this ratio change is
merely a symptom of some other change in the firm (e.g. declining profitability) then holding the latter constant may anaesthetise the gearing effect altogether.

A final comment on the classification accuracy of the financial model is in order. Whilst the number of B's is small both absolutely (a mere 17) and relative to the number of A's (there are 993 of them) 71% are classified accurately. The misclassification cost with an optimal cutoff of .16 is however small. The prediction accuracy of the model is satisfactory at 50%.

V CONCLUSIONS AND POLICY IMPLICATIONS

We have shown that it is possible to classify the bankruptcy behaviour of small UK firms rather well from financial trend data alone. We have also been able to show how the performance of models varies both with distance from the event date and span of the model. We found that the quality of the model as measured by $R^2$ and level of significance of individual coefficients tended to be higher the larger the span of the model, though the effects on model quality of increasing distance from the event date were nonmonotonic. The preferred model of the set estimated was regression 50 with reg 40 a very satisfactory alternative if data is more sparse. The coefficients of the optimal model are moreover stable and have clear economic/financial interpretations. In particular the trends of profitability and liquidity variables in the years before bankruptcy occurs constitute an fascinating cautionary tale of the small firm's eventual descent into insolvency.

Policy Conclusions: Government Agencies

The policy conclusions for government agencies flowing from the
analysis are as follows.

In a population of 1000 firms having six years' financial data about 20 will become bankrupt in any one year. One-fortieth or about 25 of these 1000 firms should be investigated. They are identified as the set of firms with bankruptcy probabilities in the model (financial trends) above 16% and predicted by the model as insolvent within one year. Within this set of twenty-five firms will be included about 70% of the actual bankrupts in the following year. About 50% or 12 of the identified firms will actually become insolvent within the period.

From the above discussion we can say that the model convincingly identifies insolvent firms in the coming year and also accurately predicts which firms of those investigated will in fact become bankrupt.

The precise value of the model to the government agency policy-maker will clearly depend on the costs of investigating 25 firms and the benefits from avoiding the insolvency of 12 such firms. However we believe the model is the best systematic guide to this kind of policy decision.

Policy Conclusions: Providers of Finance

The policy conclusions for other categories of investor, e.g. the bank or venture capitalist receiving a loan request from a small company may of course be different. He may not be interested in investigating 25 firms for signs of impending insolvency but will wish to know whether the current loan applicant is likely to go bankrupt. The model has a roughly two in three chance of telling him this correctly and is therefore potentially of considerable use in evaluating his investment prospects.
Furthermore the model predicts well the firms that will remain solvent in the next year. Therefore a venture capitalist wishing to know which firms will be good financial bets will find considerable help in the relative ranking of firms provided by the model. This ranking is in terms of probabilities of solvency and can be regarded as a measure of relative company financial performance or relative company financial risk: a higher probability indicates a firm further away from the insolvency threshold and by implication more profitable and more liquid. Thus because of the importance of profitability trends in the conditional probability measure we have derived the overall performance criterion is "highly correlated" with intuitive notions of performance used by businessmen. ¹⁰

Directions for Future Research

In a future paper it is intended to examine in the vein of SWW (1987) and Peel and Peel (1986) the effects of adding to the financial trend variables off-balance sheet items such as lags in accounts submission and audit qualification. The addition of such items can only be expected to strengthen further the predictive accuracy and practical usefulness of the present model.
References


Bathory, A The Analysis of Credit, MacMillan ,1986


Maddala, G S Limited Dependent and Qualitative Variables in Econometrics, Cambridge, 1983


Peel MJ and Peel DA, "Predicting Corporate Failure: Some Results for the UK Corporate Sector", Omega, 14(1), 1986: 5-12


1 A recent example of such deficiency is Lau(1987) who does not discuss the issue of cutoff criteria. An exception to the general rule is Martin(1973).

2 Since most studies have not considered this issue it is highly probable that their assumed cutoffs are not optimal in the above sense.

3 A legal requirement on all UK private companies is that their shares are not saleable to the general public. There are also restrictions on the transfer of shares amongst existing shareholders.

4 This formula specialises to Palepu's provided the RHS=1.

5 See Martin(1973) for this choice of loss values.

6 The analysis of this point is deferred to future paper.

7 Full-time male employees.
By an accounting identity W3 = 1 - W7. The model is therefore impossible to estimate unless one of these variables is dropped. Since most computer programs do this and the storey equations do not contain W3 we assume this has been done. The regression package we shall use later automatically deletes redundant variables. We have therefore not deleted one of the two and the results that follow must be interpreted accordingly.

For this reason it is sensible to speak of bankrupt/active observations rather than firms, since the status of a firm will in general change over time. It is worth noting that there is very little discussion of this issue in the literature, though it is as important as the questions of bias discussed in Smijewski and Palepu referred to above.

The actual number in any sample used will depend on the number of variables in the model being estimated. In general the larger the number of variables the smaller the effective size due to missing data on some or all variables.

Here are several problems with classification accuracy (defined as the proportion of bankrupts (actives) correctly classified as such) as a criterion of the success of a model. The most obvious are however (i) probabilities are continuous whereas the event studied is categorical; hence classification accuracy may fluctuate wildly from sample to sample unless the estimated probabilities are well separated from the cutoff point; (ii) the measure ignores associated Type I and Type II error costs so that models with large and small prediction error rates are rated as equivalent.
Complete here means that there is a nonmissing value for all the variables in the observation.

This is of course not a random sample of small firms but may be thought of as a substantially less nonrandom sample of firms with (up to) six years' complete data on sixteen variables. Zmijewski (1984) the only writer to have examined the impact of excluding observations with missing data from bankruptcy analysis found that in practice it had only a small effect on coefficient bias. Whilst we do not attempt to estimate here the effect of missing data Zmijewski's results give us some hope that our own estimates will not differ too far from the true values. Furthermore there is a genuine problem of misspecification present in the choice of the variable to estimate the probability of missing data. About the only meaningful variable for this purpose in the SKWW database would seem to be the age of the firm. However this is actually missing for about 300 of the 636 firms! A detailed discussion of this issue is left to another paper.

To enhance interpretation of the variables of the model we have standardised all variables to zero mean and unit variance. This allows regression coefficient ratios to represent marginal rates of substitution between corresponding variables along probability isoquants when evaluated at variable means.