UNIVERSITY OF EDINBURGH
Business School

CREDIT RESEARCH CENTRE

Chinese Companies Distress Prediction: an application of DEA

Zhiyong Li, Jonathan Crook, Galina Andreeva

25.08.2011

UNIVERSITY OF EDINBURGH Business School

Introduction

- Financial information

Accounting data and financial ratios in statements (since Beaver, 1966)

Altman (1968): Multiple Discriminant Analysis, Z-score

$Z = .012X_1 + .014X_2 + .033X_3 + .006X_4 + .999X_5$

where X_1 = Working capital/Total assets

X_2 = Retained Earnings/Total assets

X_3 = Earnings before interest and taxes/Total assets

X_4 = Market value equity/Book value of total debt

X_5 = Sales/Total assets

- Corporate performance

Xu and Wang (2009): in Support Vector Machines (SVMs) and Multiple Discriminant Analysis (MDA)

Yeh et al. (2010):in the integrated Rough Set Theory (RST) with SVM

Paradi et al. (2004): the worst practice DEA

Objectives

To predict corporate failures by

- Corporate performance measures

Variable Returns to Scale (VRS) assumption

Return to scale levels

Cross-sectional, panel, and survival models

- Logistic Regression

with efficiencies

without efficiencies

- Comparison with Altman’s model

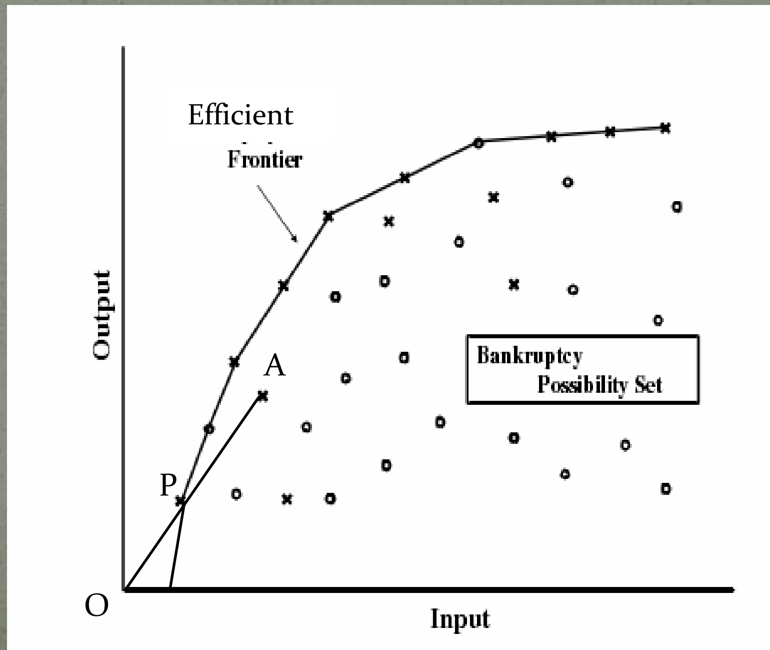
DEA

- Performance measurement

Performance is commonly measured by $\frac{\text{Output}}{\text{Input}}$, called ‘efficiency’ or ‘productivity’

- Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a method to measure ‘relative efficiency’ of Decision Making Units (DMUs). (Charnes, Cooper & Rhodes, 1978)



$$\text{relative efficiency} = \frac{OP}{OA}$$

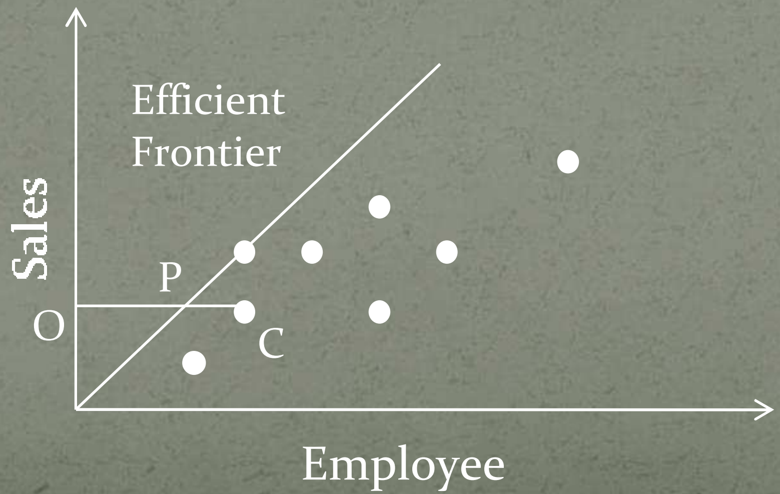
UNIVERSITY OF EDINBURGH
Business School

DEA

Some simple examples

- One input and one output

Store	A	B	C	D	E	F	G	H
Employee	2	3	3	4	5	5	6	8
Sale	1	3	2	3	4	2	3	5
Efficiency	0.5	1	0.67	0.75	0.8	0.4	0.5	0.625

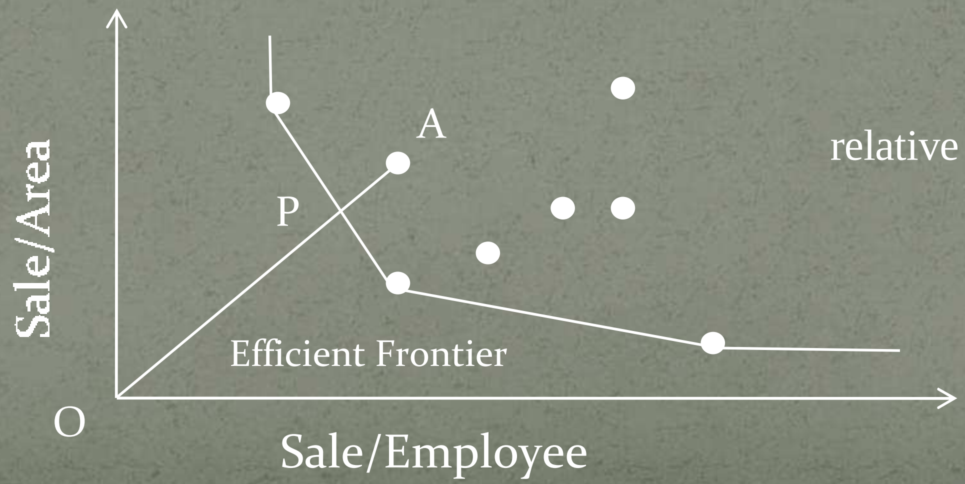


$$\text{relative efficiency} = \frac{OP}{OC}$$

DEA

- Two inputs and one output

Store	A	B	C	D	E	F	G	H
Employee	4	7	8	4	2	5	6	5
Area	3	3	1	2	4	2	4	2
Sale	1	1	1	1	1	1	1	1



$$\text{relative efficiency} = \frac{OP}{OA}$$

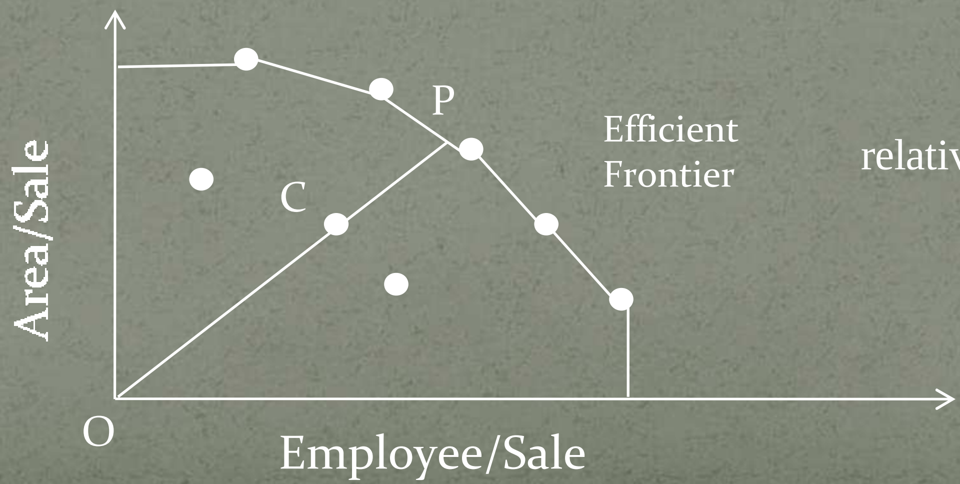
UNIVERSITY OF EDINBURGH Business School

Credit Research Centre

DEA

- One input and two outputs

Store	A	B	C	D	E	F	G	H
Employee	2	3	3	4	5	5	6	8
Customer	1	2	3	4	4	5	5	6
Sale	1	3	2	3	4	2	3	5



$$\text{relative efficiency} = \frac{OC}{OP}$$

DEA

- The basic CCR model

CCR is named by Charnes, Cooper & Rhodes (1978)

n DMUs

Weights of inputs

m inputs

Weights of outputs

s outputs

		1	2	3	…	j	…	n		
v_1	1	x_{11}	x_{12}	x_{13}	…	x_{1j}	…	x_{1n}		
v_2	2	x_{21}	x_{22}	x_{23}	…	x_{2j}	…	x_{2n}		
	.	.	.	.	.	.	…	.		
v_i	.	.	.	.	.	x_{ij}	…	.		
	.	.	.	.	.	.	…	.		
v_m	m	x_{m1}	x_{m2}	x_{m3}	…	x_{mj}	…	x_{mn}		
		y_{11}	y_{12}	y_{13}	…	y_{1j}	…	y_{1n}	1	u_1
		y_{21}	y_{22}	y_{23}	…	y_{2j}	…	y_{2n}	2	u_2
		.	.	.	.	.	…	.	.	
		.	.	.	.	y_{rj}	…	.	.	u_r
		.	.	.	.	.	…	.	.	
		y_{s1}	y_{s2}	y_{s3}	…	y_{sj}	…	y_{sn}	s	u_s

UNIVERSITY OF EDINBURGH
Business School

DEA

- The basic CCR model

For each DMU_j, the efficiency is measured by:

$$\theta_j = \frac{u^T y_i}{v^T x_j} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}, \; j = 1, 2, \ldots, n$$

Let the DMU_j, to be evaluated on any trial be designated as DMU_o where *o ranges over 1, 2, …, n.* We have the fractional programming problem to solve the weights of inputs and outputs.

$$(FP_o) \quad \max \quad \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \cdots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \cdots + v_m x_{mo}}$$

subject to

$$\frac{u_1 y_{1j} + u_2 y_{2j} + \cdots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_{2j} + \cdots + v_m x_{mj}} \le 1 \; (j = 1, \cdots, n)$$

$$v_1, v_2, \cdots, v_m \ge 0$$

$$u_1, u_2, \cdots, u_m \ge 0$$

DEA

- Return to Scale

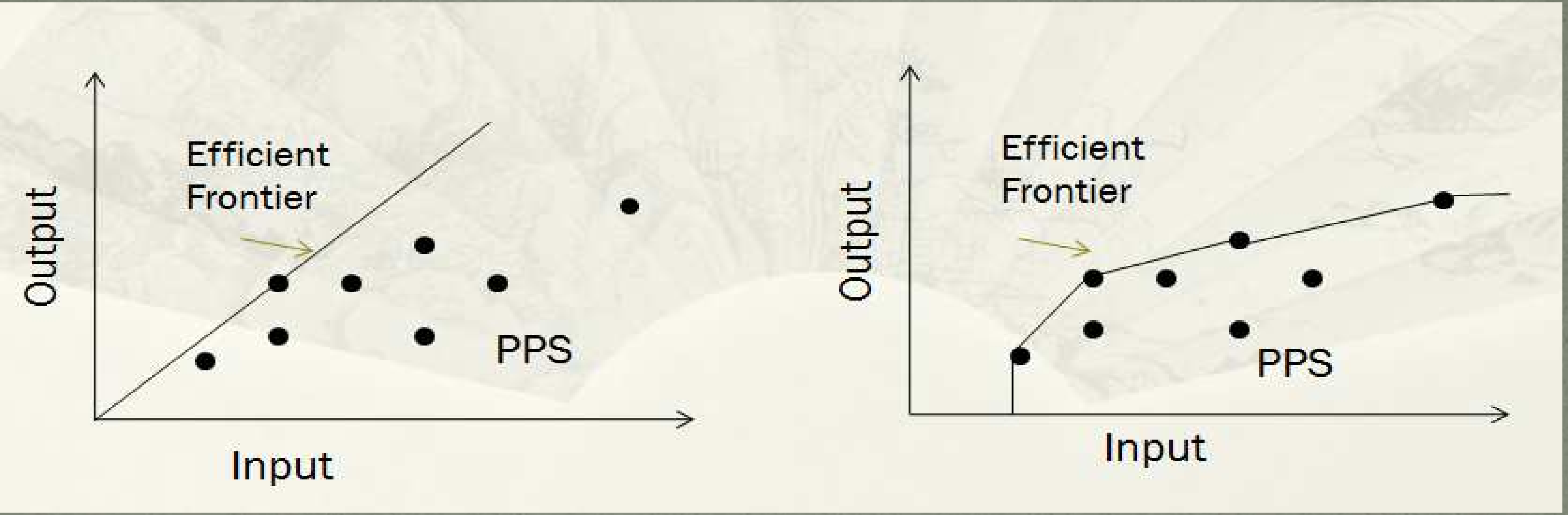
Returns to Scale (RTS) is the term to describe what happens as the scale of production increases when all inputs and outputs are variables.

Constant Returns to Scale (CRS): when the relative change in output is the same compared to the relative change in input

Variable Returns to Scale (VRS): If the proportional increase in output is larger (smaller) than the proportional increase of input, it is increasing (decreasing) returns to scale.

DEA

- BCC model (VRS is assumed, Banker et al., 1984)



$$
\begin{aligned}
(BCC) \quad \min \quad & \theta_B \\
\text{s.t.} \quad & \theta_B x_o - X\lambda \geq 0 \\
& Y\lambda \geq y_o \\
& e\lambda = 1 \\
& \lambda \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\max \quad & z = uy - u_0 \\
\text{s.t.} \quad & vx_o = 1 \\
& -vX + uY - u_0 e \leq 0 \\
& v \geq 0, u \geq 0, u_0 \text{ free in sign}
\end{aligned}
$$

DEA

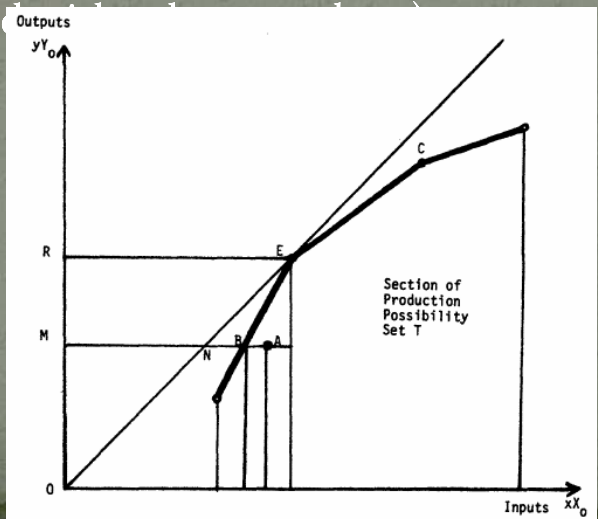
- Four predictors calculated by DEA

Pure Technical Efficiency: the potential productivity which can be achieved by optimization of inputs and outputs, from the technical point of view (the ability to utilize input efficiently).

Scale efficiency: the potential productivity gain from achieving optimal size of a firm.

Overall Technical Efficiency: simply the product of Pure Technical Efficiency and Scale efficiency. (Banker, et al. 1984)

Return to Scale Estimation: an indicator to denote on which stage the company is operating, decreasing, increasing or constant, within the same industrial sector (compare



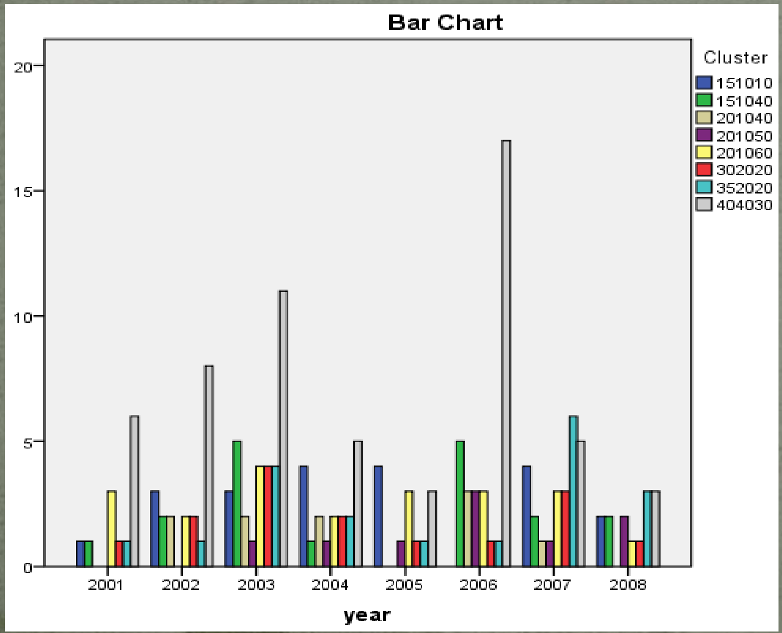
$$\text{Pure Technical Efficiency}=\frac{MB}{MA}$$

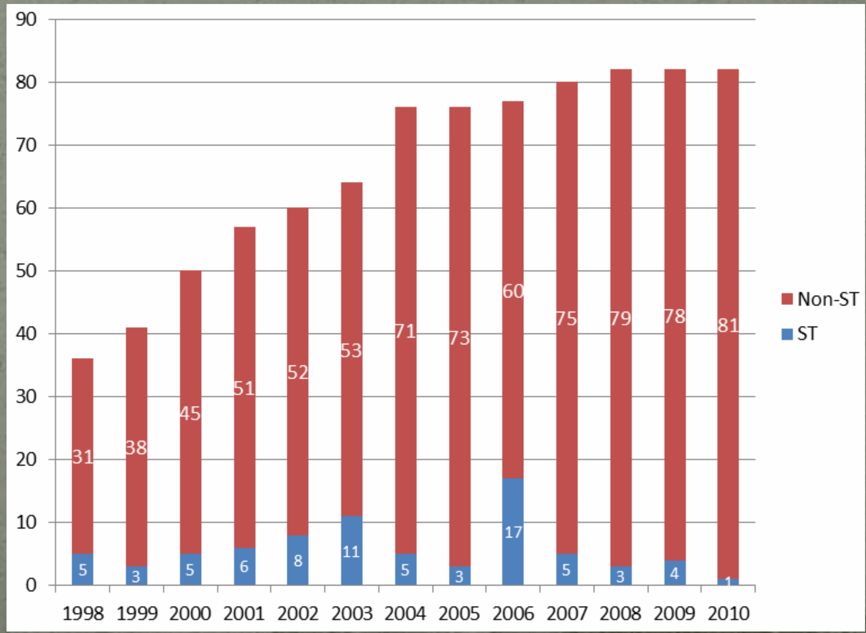
$$\text{Scale Efficiency}=\frac{MN}{MB}$$

$$\text{Overall Technical Efficiency}=\frac{MN}{MA}$$

Data

- all Chinese listed companies (over 2,000) from 1991 to present.
- Financial distress indicator: Special Treatment (defined by China Securities Regulatory Commission)
- Since DEA requires homogeneity (the same productivity function in the sample), the industry sector Real Estate is found to be the one with most BAD cases.





UNIVERSITY OF EDINBURGH
Business School

Data

- Financial ratios

Ratio groups	In database (89)	After deleting (52)
Indicator per share	15	11
Profitability	20	15
profit composition	5	0
Capital composition	9	8
Liquidity	16	11
Operation capacity	8	3
Cash flow	4	2
Growth rates	12	2

UNIVERSITY OF EDINBURGH
Business School

Data

- DEA inputs and outputs

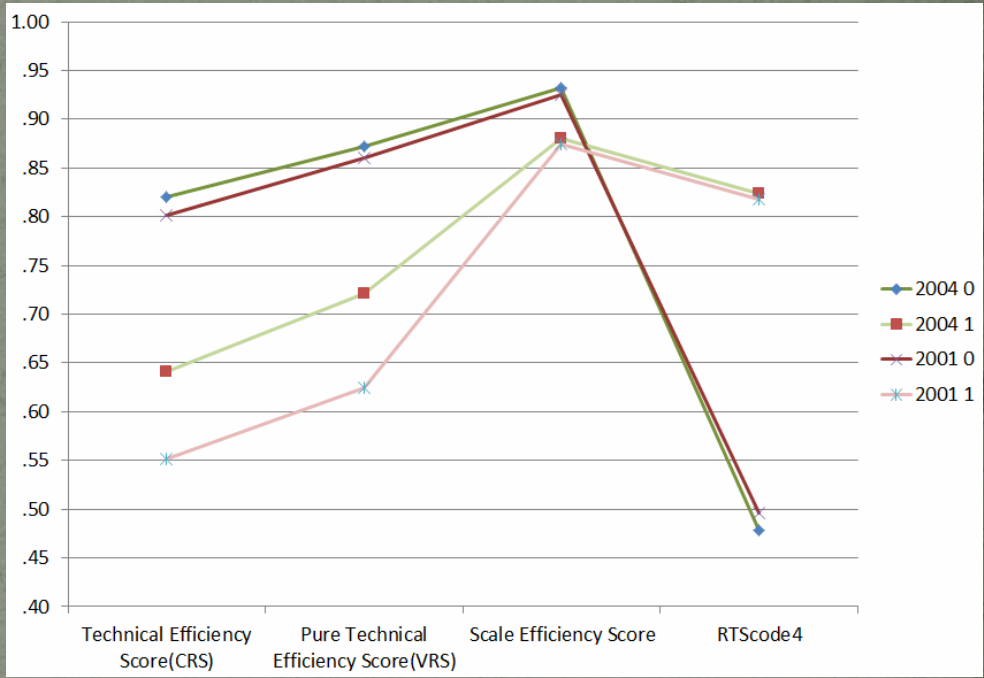
Year 2001 (N=130)	totalsales (m)	totalcost (m)	totalprofits (m)	totalassets (m)	totaldebts (m)	sharecapital (m)	cashaccrued (m)	staff
Mean	516	489	39	1540	792	285	47	1150
Median	315	302	28	1090	499	219	4	729
Std. Deviation	672	620	123	1470	940	227	158	1570
Minimum	0	14	-538	59	6	54	-330	15
Maximum	4460	4160	502	9690	7380	1870	819	13300
Kurtosis	13.028	13.205	5.923	9.012	18.98	18.103	6.16	28.093
Skewness	3.253	3.266	-0.479	2.555	3.563	3.257	1.979	4.246

Year 2004 (N=134)	totalsales (m)	totalcost (m)	totalprofits (m)	totalassets (m)	totaldebts (m)	sharecapital (m)	cashaccrued (m)	staff
Mean	732	703	27	2020	1200	328	27	945
Median	438	466	27	1360	789	250	-2	479
Std. Deviation	939	821	229	2100	1280	298	304	1630
Minimum	0	0	-954	120	5	54	-600	24
Maximum	7670	6420	1260	15500	9230	2270	2160	13600
Kurtosis	13.028	13.205	5.923	9.012	18.98	18.103	6.16	28.093
Skewness	3.253	3.266	-0.479	2.555	3.563	3.257	1.979	4.246

Results

- DEA results

mean score		Technical Efficiency Score(CRS)	Pure Technical Efficiency Score(VRS)	Scale Efficiency Score	RTScode
2004	0	.82	.87	.93	.48
	1	**.64**	**.72**	**.88**	**.82**
	all	.80	.85	.93	.52
2001	0	.80	.86	.92	.50
	1	**.55**	**.62**	**.87**	**.82**
	all	.78	.84	.92	.52



Results

- Training sample

Independent variables: 2001

Distress indicator: 2003 (Good/Bad: 116/11)

Model 1: Stepwise Logistic, ratios only

Model 2: Stepwise Logistic, ratios & efficiencies

Model 3: Enter Logistic, significant ratios in 1 & 2 and efficiencies

Variables in the Equation

		B	S.E.	Wald	df	Sig.
Step 1a	TechnicalEfficiencyScoreCRS	178.103	98.039	3.300	1	.069
	PureTechnicalEfficiencyScoreVRS	-252.570	126.107	4.011	1	.045
	ScaleEfficiencyScore	-147.702	74.025	3.981	1	.046
	RTScode	-1.538	.911	2.847	1	.092
	EPSatreportdatetotalshare	-3.783	2.122	3.177	1	.075
	equitymultiplier	-2.982	1.936	2.372	1	.124
	netassetsattributabletoequity／investedcapital	-.090	.078	1.317	1	.251
	totalexpensestosalesratio	-.225	.098	5.291	1	.021
	currentassetsturnover	8.716	3.893	5.013	1	.025
	Constant	209.707	95.247	4.848	1	.028

UNIVERSITY OF EDINBURGH
Business School

Results

CREDIT RESEARCH CENTRE

- Test sample

Independent variables: 2004

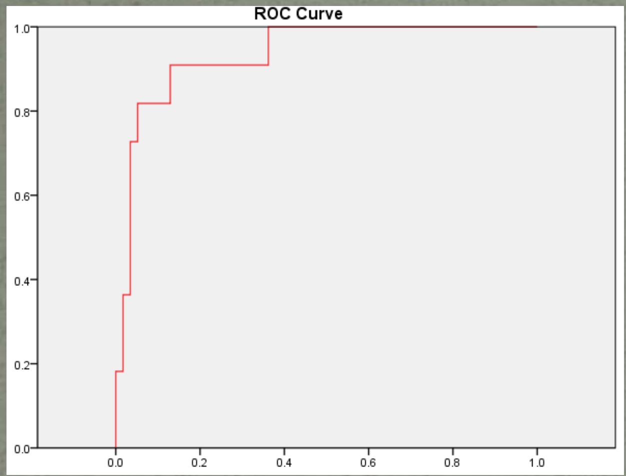
Distress indicator: 2006 (Good/Bad: 113/17)

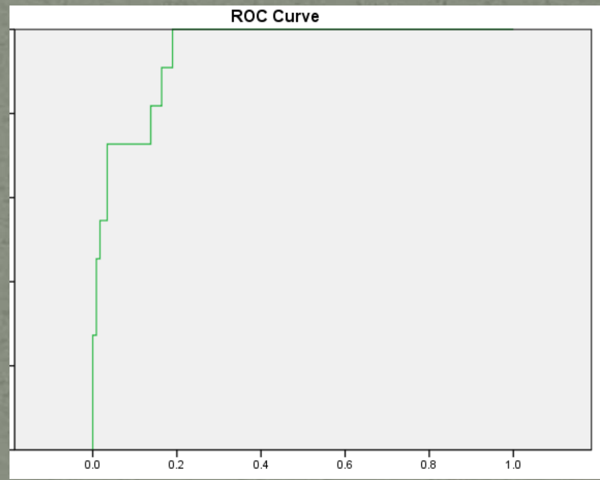
Model 4: Enter Logistic, variables and their coefficients in Model 3.

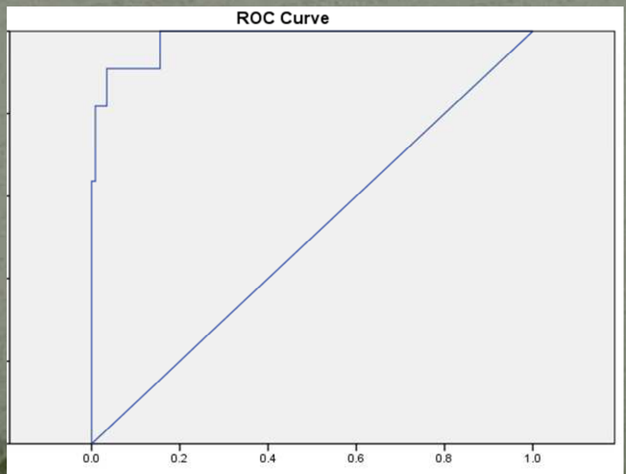
	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square	AUROC
Model 1	42.800	.223	.501	.935
Model 2	38.459	.249	.559	.946
Model 3	24.148	.329	.739	.981
Model 4				.679

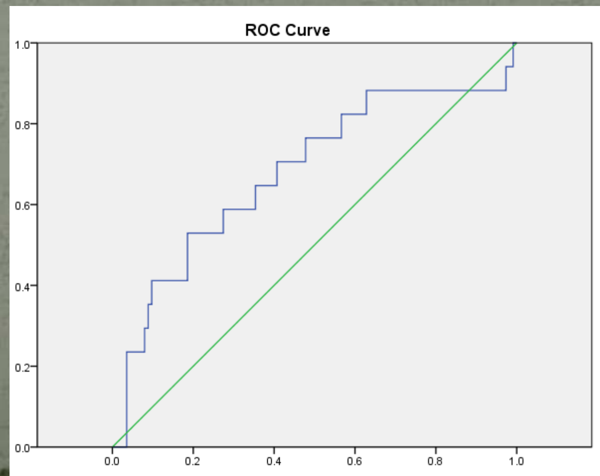
Results

- ROC curve









Results

- Compare with Altman's Z-score

Model 4				
		Predicted		
		ST06		Percentage Correct
Observed		0	1	
ST06	0	101	12	89%
	1	10	7	41%
Overall Percentage				83.08%

Z-score				
		Predicted		
		ST06		Percentage Correct
Observed		0	1	
ST06	0	101	12	89%
	1	12	5	29.4%
Overall Percentage				82.3%

Conclusion

- Conclusion

Corporate performance measurements (efficiencies) improve predictive accuracy when running with financial ratios in Logistic Regression

The more efficient (in the way of optimal operation and scale), the less probability a company goes distressed.

Increasing return to scale is associated to financial distress.

- Future work

Malmquist model

Panel pooled data regression over 10 years

Survival model

Thank you!