

Financial and Economic Determinants of Firm Default

Giulio Bottazzi Marco Grazzi Angelo Secchi Federico Tamagni

LEM, Scuola Superiore Sant'Anna, Pisa

HEC Management School of the University of Liège
20 Octobre 2009

Outline

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Motivation & Background

Default Determinants

- ① Financial economics: distress prediction is a short term issue, determined by financial conditions (exposure, leverage, etc...) only
- ② Industrial Economics: how / why firms grow, survive or exit, based on economic/industrial characteristics of (heterogeneous) firms

BUT

- How does economic dynamics (growth, selection, creative destruction) affects financial default ? Is it default only financial ?
- Does short term financial performance perfectly embed economic/industrial performance of a firm ?

⇒ Complement economic (productivity, profitability, size, growth) and financial characteristics, over different time distance to default

Economic exit

Search for profit drives the entry/exit dynamics.

Symmetric framework:

Higher efficiency \rightarrow Profit opportunities \rightarrow Entry

Lower efficiency \rightarrow (Expected) Losses \rightarrow Exit.

Firms grow, survival or exit, based on economic/industrial characteristics:
underlying assumption in industrial dynamics literature (Lucas, Winter, Jovanovic, Ericson-Pakes, ...).

Exit is an economic decision

Financial default

Default is often a (not anticipated) traumatic event.

It is determined by financial conditions (exposure, leverage, etc...) leading to insolvency.

Its prediction is a short term issue.

Default is a traumatic occurrence

When default occurs

Definitions according to Basel II - International Convergence of Capital Measurement and Capital Standard

A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place

- The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realising security (if held)
- The obligor is past due more than 90 days [...]

Bridging economic and financial vision

General question: has default only financial causes or is it related to the economic characteristics of a firm?

Does short term financial performance perfectly embodies the long term economic/industrial performance of a firm?

What about market imperfections and frictions?

In practice: Analyze economic (productivity, profitability, size, ...) and financial characteristics in the years before default occurs.

Methodology: from exploratory data analysis to parametric (probit regressions)

Sources

Database with firm economic and financial variables and default events.

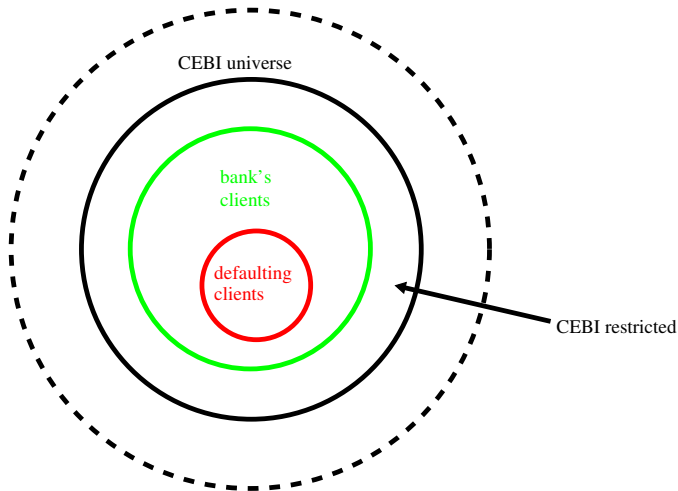
=

Default events provided by a large Italian banking group: ~ 150 manufacturing firms in 2003 or 2004.

+

Financial statements and balance sheets from CeBi database, virtually the universe of *limited liability* firms. (~ 40000)

Harmonization



Solution: Restrict the definition of the universe

Harmonization procedure

- Equipopulated samples in each year: 1998-2003.
- Comparability with the subset of defaulting firms: annual sales above 1 ML €.
- Minimum scale of operations: above 1 employee (avoid self-employment).

Variables selection

Economic variables

- SIZE = Sales (Annual Revenues)
- GROWTH = $\Delta \ln$ Sales
- PROF = ROS (Operating Margins/Sales)
- PROD = Value Added per Employee

Financial variables

- IE/S = Interest Expenses, scaled by size
- FD/S = Financial Debt-to-Sales ratio
- Leverage = Debt/Equity



Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Kernel estimates

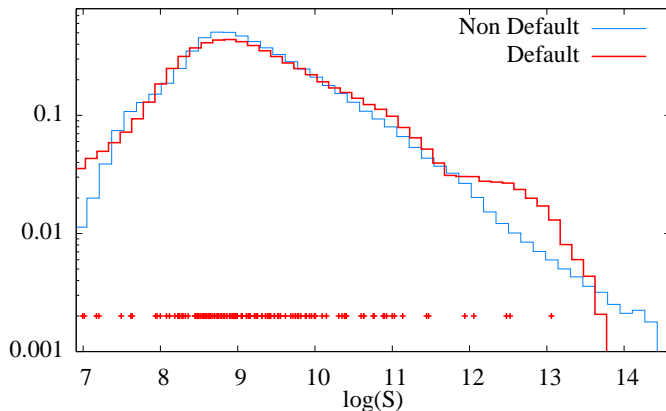
Estimate density $\hat{f}(x)$ from a set of observations $\{x_1, \dots, x_N\}$

$$\hat{f}(x) = \frac{1}{N} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right)$$

depends (weakly) on kernel K and (strongly) on bandwidth h .

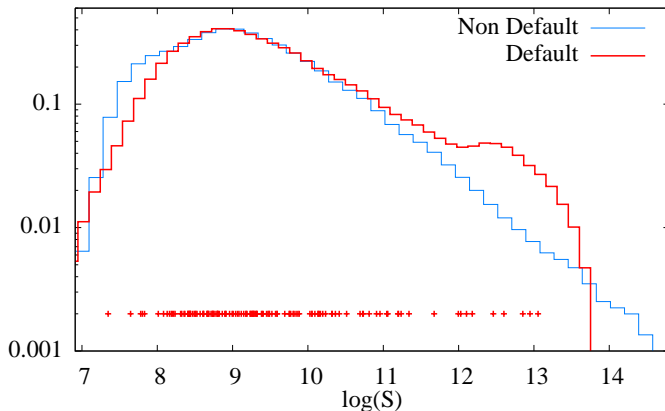
Compare distributional properties of variables for default (D) and non-default (ND) firms.

Size empirical densities, 1998



Similarly distributed

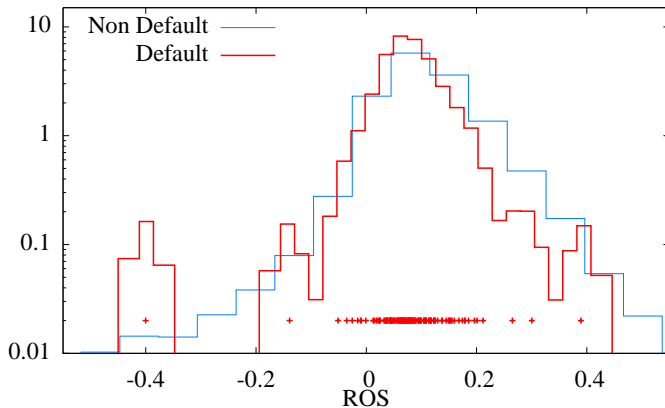
Size empirical densities, 2002



No time effect

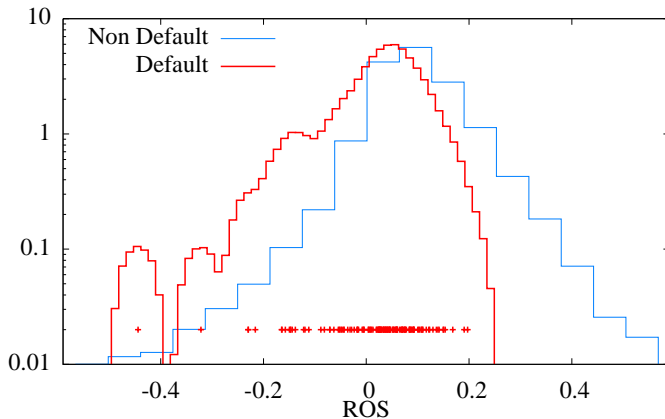


Profitability empirical densities, 1998



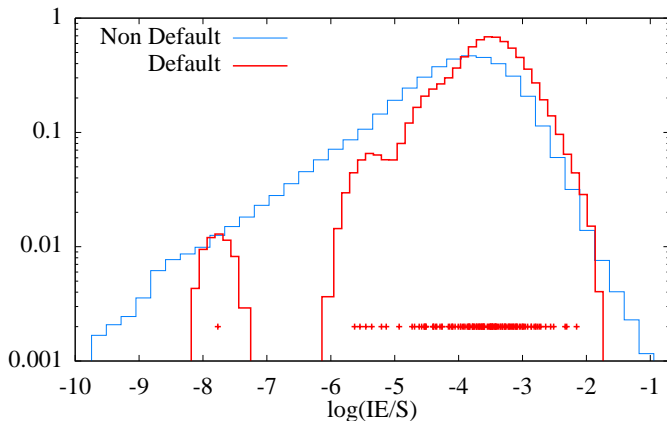
Similarly distributed

Profitability empirical densities, 2002



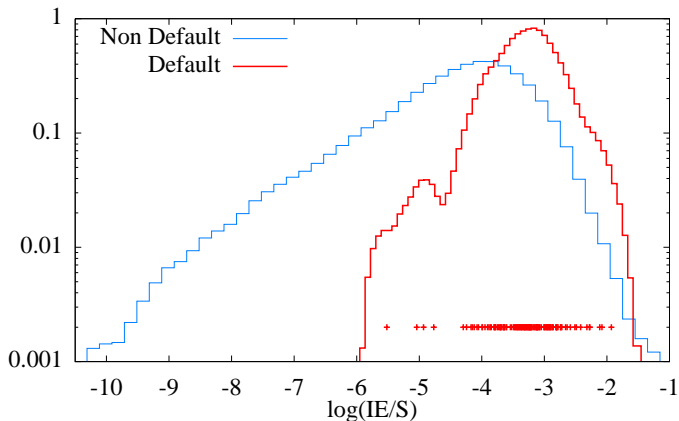
Progressive tilt in time

IE/S empirical densities in 1998



Distributional differences

IE/S empirical densities in 2000



Progressive shift in time



Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- **Non-parametric analysis**
- Probit regression

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Stochastic inequality

Problems with D and ND samples:

- they differ in size
- distributions differ in shape (heteroskedasticity)
- distributions are asymmetric (weakness of central statistics)

Rank according to stochastic (in) equality

$$\mathbf{X}_{ND} \text{ larger than } \mathbf{X}_D \text{ iff } \text{Prob}\{X_{ND} > X_D\} > 1/2$$

Using Fligner and Policello (1981) test the equivalent

$$H_0 : \int dF_D F_{ND} = \frac{1}{2} \quad \text{vs} \quad H_1 : \int dF_D F_{ND} \neq \frac{1}{2} .$$



FP test: financial variable

		Test of Stochastic Equality				
Variable	Test	1998	1999	2000	2001	2002
IE/S	FP stat	7.510	9.269	13.019	17.903	24.069
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000
LEV	FP stat	8.029	10.483	12.066	13.520	15.190
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000
FD/S	FP stat	7.490	10.480	14.387	16.037	17.229
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000

Statistic and associated p -value. 1% confidence level in bold.

FP test: economic variables

Test of Stochastic Equality						
Variable	Test	1998	1999	2000	2001	2002
SIZE	FP stat	0.364	1.555	3.988	3.466	2.426
	p-value	0.716	0.120	0.000	0.000	0.015
GROWTH	FP stat		0.905	-0.618	-1.133	-3.927
	p-value		0.365	0.536	0.257	0.000
PROF	FP stat	-4.609	-7.169	-7.186	-7.466	-11.176
	p-value	0.000	0.000	0.000	0.000	0.000
PROD	FP stat	-5.310	-7.156	-7.167	-6.842	-8.855
	p-value	0.000	0.000	0.000	0.000	0.000

Statistic and associated p -value. 1% confidence level in bold.

Non-parametric results

- 1 Financial Vars: as expected Non-Defaulting firms perform better
- 2 Economic Vars:
 - ▶ PROD and PROF: as expected, Non-Defaulting firms perform better
 - ▶ SIZE and GROWTH: not completely expected, the two groups are quite comparable

Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- **Probit regression**

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Different specifications

Probit model:

$$P(Y_T = 1 | X_t) = \Phi \left(\beta_0 + \sum_j \beta_j X_j \right)$$

Model specifications:

- Financial Variables: $X \in \{\text{IE/S, LEV, FD/S}\}$
- Financial+Economic: $X \in \{\dots, \text{SIZE, PROD, PROF, GROWTH,}\}$

Time structure:

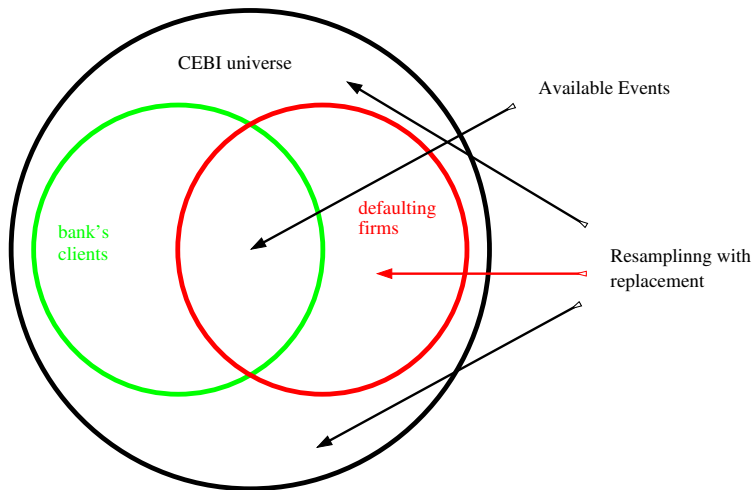
- Year by year regression: $X = X_t \quad t \in \{1, \dots, T\}$

Z-score of variables to allow comparability among coefficients

Sectors	Number of firms	Number of defaults	Default rate Sample	Default rate population
15 - Food products, beverages	2008	9	0.0045	0.0302
17 - Manufacture of textiles	1544	14	0.0091	0.0474
18 - Wearing apparel	673	9	0.0134	0.0511
19 - Leather, footwear	762	15	0.0197	0.0375
20 - Manufacture of wood	396	2	0.0051	0.0249
21 - Pulp & paper products	540	2	0.0037	0.0298
22 - Publishing, printing	749	6	0.0080	0.0377
24 - Chemical products	1122	2	0.0018	0.0383
25 - Rubber, plastic products	1176	6	0.0051	0.0338
26 - Other non-metallic	1142	5	0.0044	0.0309
27 - Manufact of basic metals	672	2	0.0030	0.0378
28 - Metal products	2473	14	0.0057	0.0280
29 - Machinery and equipment	2916	23	0.0079	0.0352
continues in the paper ...				



Bootstrap approach



Solution: Bootstrap re-sampling with replacement.

Re-sampling procedure

Procedure:

- “Reproduce”, in each replication, the default/non default ratio at 2-Digit industry level (data available from Chambers of Commerce).
- Defaulting firms fixed, and re-sampling of Non-Defaulting only
- Robust statistical significance via bootstrap percentiles

Year by year: estimates and significance

	1999	2000	2001	2002	1999	2000	2001	2002
IE/S	0.0051*	0.0056*	0.0073*	0.0139*	0.0054*	0.0055*	0.0071*	0.0130*
LEV	0.0039	0.0032	0.0026	0.0063*	0.0026	0.0020	0.0014	0.0057*
FD/S	0.0067*	0.0080	0.0064	0.0053	0.0049	0.0072	0.0048	0.0034
SIZE					0.0060*	0.0076*	0.0099*	0.0097*
PROD					-0.0128*	-0.0093*	-0.0072*	0.0023
PROF					-0.0012	-0.0015	-0.0040	-0.0068*
GROWTH					0.0062*	0.0009	-0.0022	-0.0045

* Significant at 1% level.

Year by year: performance and comparisons

	1999	2000	2001	2002	1999	2000	2001	2002
Panel B: Model performance								
Brier Score	0.0333	0.0336	0.0336	0.0332	0.0330	0.0334	0.0335	0.0330
Threshold	0.0313	0.0321	0.0320	0.0300	0.0333	0.0323	0.0324	0.0353
Type I error	30.3500	31.9300	23.1650	18.2950	36.5700	29.0950	27.9700	25.3400
Type II error	1541.61	1425.80	1454.20	1137.66	1322.49	1466.08	1425.74	886.74
% Correct default	0.7629	0.7635	0.8333	0.8603	0.7143	0.7845	0.7988	0.8066
% Correct non default	0.5712	0.6209	0.6250	0.6880	0.6321	0.6102	0.6324	0.7568
Panel C: Comparison with financial variables only								
Threshold					0.0313	0.0321	0.0320	0.0300
Type I error					29.1150	28.5450	26.7150	17.4650
Type II error					1542.13	1483.18	1459.94	1152.60
% Correct default					0.7725	0.7886	0.8078	0.8667
% Correct non default					0.5710	0.6056	0.6235	0.6839



Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

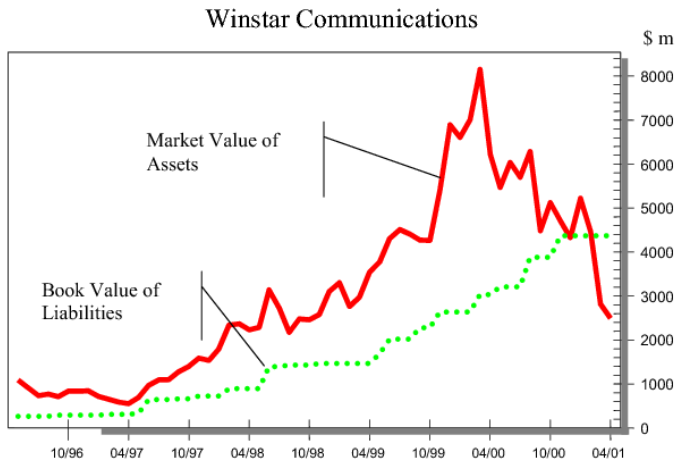
- **Distance to Default**
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Dinstance of assets to liabilities



Evolution of asset value and book liabilities (src. Moody's).

Motivation

Equity value modeled as a call option on firm's assets and strike price equal to firm's liability.

From present equity price infer the probability to exercise the option, that is $\text{assets} > \text{liabilities}$.

Grounded in theoretical financial literature (Merton, 1979).

Particularly suitable for publicly traded companies

But this is not our case!.

“book DD” inclusion

We start from a “naive DD” (Bharath and Shumway, 2008)

$$\text{naive DD} = \frac{\ln[(E + F)/F] + (r_{t-1} - 0.5 \text{ naive } \sigma_V^2)T}{\text{naive } \sigma_V \sqrt{T}} ,$$

E is equity, F is debt, σ_V assets volatility, r price return and T time to maturity.

- Replace Equity E with Book Equity $BE = TA/LEV$.
- Define total debt as $TA - BE$.
- Use growth rate of BE for r .
- Volatility of firm's value as weighted average

$$\sigma_V = \frac{BE}{BE + D} \sigma_{BE} + \frac{D}{BE + D} \sigma_D .$$

Year by year with DD: estimates and significance

	1999	2000	2001	2002	1999	2000	2001	2002
IE/S					0.0074*	0.0066*	0.0086*	0.0090*
LEV					0.0005	0.0015	0.0002	0.0029
FD/S					0.0024	0.0053	0.0021	0.0026
ln SIZE					0.0066*	0.0076*	0.0075*	0.0072*
PROD					-0.0097*	-0.0067*	-0.0041	-0.0010
PROF					0.0013	-0.0005	-0.0024	-0.0049*
GROWTH					0.0053*	0.0004	-0.0019	-0.0020
Book DD	-0.0210*	-0.0204*	-0.0898*	-0.0816*	-0.0164*	-0.0147*	-0.1095*	-0.1064*

* Significant at 1% level.

Year by year with DD: performance and comparisons

	1999	2000	2001	2002	1999	2000	2001	2002
Panel B: Model performance								
Brier Score	0.0332	0.0335	0.0335	0.0335	0.0329	0.0334	0.0333	0.0331
Threshold	0.0374	0.0385	0.0364	0.0362	0.0343	0.0315	0.0345	0.0361
Type I error	43.9900	51.6650	52.2200	49.2350	32.6450	22.6450	30.2400	29.2500
Type II error	1224.45	1180.96	1305.32	1358.47	1159.70	1431.33	1131.60	901.65
% Correct default	0.6364	0.5964	0.6014	0.6242	0.7302	0.8231	0.7692	0.7767
% Correct non default	0.6386	0.6677	0.6420	0.6274	0.6577	0.5973	0.6896	0.7527
Panel C: Comparisons of prediction performance against DD only								
Threshold					0.0374	0.0385	0.0364	0.0362
Type I error					41.6750	46.8650	34.7800	29.5750
Type II error					951.95	974.07	1020.30	892.44
% Correct default					0.6556	0.6339	0.7345	0.7742
% Correct non default					0.7190	0.7259	0.7202	0.7552



Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

- Distance to Default
- **Risk rating**

5 Conclusions

- Variables relevance
- Probit analysis



Motivation

Credit ratings

- include a range of dimensions which we could have omitted
- provide a short term (one year) forecast of default, useful to test the observed short time relevance of economic variables

Use of CeBi Rating Index

- Former official rating of Bank of Italy
- Issuer rating index
- CeBi is Italian member of ECCBSO (Eur. Comm. of Central Balance Sheet Data Offices)
- Available for every single firm included in the database

Risk rating inclusion

Rating originally provided in 1 – 9 range but not cardinal.

Assigned to 3 rating classes: High, Medium and Low Risk. Dummy variables introduced as deviation from Low Risk.

Transition Probabilities



Year by year with CR: estimates and significance

	1999	2000	2001	2002		1999	2000	2001	2002
IE/S						0.0039*	0.0041*	0.0049	0.0088*
LEV						-0.0001	-0.0013	-0.0015	0.0028*
FD/S						0.0044	0.0059	0.0026	0.0019
ln SIZE						0.0061*	0.0072*	0.0090*	0.0086*
PROD						-0.0111*	-0.0075*	-0.0059*	-0.0011
PROF						-0.0004	-0.0015	-0.0021	-0.0037*
GROWTH						0.0054*	-0.0000	-0.0038*	-0.0036*
LOW	-0.0301*	-0.0076	-0.0323*	-0.1206*	0.0001	0.0186*	-0.0047	-0.0428*	
MID	0.0198*	0.0614*	0.0414*	-0.0072	0.0392*	0.1134*	0.0639*	0.0068	

* Significant at 1% level.

Year by year with CR: performance and comparisons

	1999	2000	2001	2002	1999	2000	2001	2002
Panel B: Model performance								
Brier Score	0.0328	0.0327	0.0322	0.0319	0.0325	0.0326	0.0320	0.0316
Threshold	0.0258	0.0284	0.0193	0.0163	0.0290	0.0254	0.0239	0.0270
Type I error	78.0000	70.7850	61.0700	48.7550	36.6350	30.0450	29.8300	25.8250
Type II error	576.19	616.35	709.39	633.72	1248.93	1334.55	1230.21	901.15
% Correct default	0.3906	0.4757	0.5606	0.6278	0.7138	0.7774	0.7854	0.8029
% Correct non default	0.8397	0.8361	0.8171	0.8262	0.6526	0.6452	0.6828	0.7528
Panel C: Comparisons of prediction performance against the "Rating only" model of the same year								
Threshold					0.0258	0.0284	0.0193	0.0163
Type I error					25.2550	39.5450	16.2850	11.8150
Type II error					1612.86	1111.91	1722.58	1507.69
% Correct default					0.8027	0.7071	0.8828	0.9098
% Correct non default					0.5514	0.7044	0.5558	0.5865



Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- **Variables relevance**
- Probit analysis



Relevance of economic variables

- Financial Vars: cost of debt (IE/S) is the most relevant factor
- Economic Vars:
 - Negative effect of PROD and PROF
 - Positive effect of SIZE and GROWTH
- Time effects:
 - Economic variables still significant in the short run
 - Trade-off PROF vs. IE/S in 2002, 1-year before default

Problem

1 Introduction

2 Data

3 Statistical analysis

- Graphical analysis
- Non-parametric analysis
- Probit regression

4 Robustness and extensions

- Distance to Default
- Risk rating

5 Conclusions

- Variables relevance
- Probit analysis



Bootstrap probit regressions: sum up

What is more:

- Sectoral specificities (2 digit) do not matter
- Including Economic Vars increases prediction accuracy (Type I errors)
- Results robust to the inclusion of credit rating and naive Distance-to-Default indicator

Goodness of fit

Brier's score: $\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{P}_i)^2$ Y_i observed, \hat{P}_i estimated .

Define a model dependent optimal threshold:

$$\tau^* = \arg \min_{\tau} \left(\frac{1}{N_0} \sum_{i \in ND} \Theta(P_i - \tau) + \frac{1}{N_1} \sum_{i \in D} \Theta(\tau - P_i) \right) ,$$

Type I error if $\hat{P}_i < \tau^*$ but $Y_i = 1$.

Type II error if $\hat{P}_i > \tau^*$ but $Y_i = 0$.

Models comparison

Assume model A has optimal threshold τ_A^* and Model B has optimal threshold τ_B^* .

Model B “beats” model A for Type I (II) errors if it gives less Type I (II) errors with τ_A^* .

Ratings transition matrix

		2003			
		Low	Mid	High	Default
1998	Low	0.8947	0.0825	0.0176	0.0052
	Mid	0.5406	0.3647	0.0743	0.0204
	High	0.4589	0.3290	0.1991	0.0130
2002	Low	0.9308	0.0583	0.0077	0.0032
	Mid	0.3377	0.5444	0.0987	0.0192
	High	0.1502	0.3519	0.4742	0.0236

