The Discriminative Power of Rating Functions

GfKl 2006

Berlin, March 8th, 2006

C. Beinert / P. Reichling / B. Vogt

Otto-von-Guericke-University Magdeburg

Agenda

1 Introduction

1.1 Motivation

- Estimating creditworthiness represents a vital success factor in the loan business of each bank
- Validation of rating systems is necessary for IRBA application (at least in Germany)
- Our aim: presentation of practicable measures + discussion of limits to their application
- As well as: introduction of the critera of 1st and 2nd order stochastic dominance to evaluate the discriminative power of rating functions

1 Introduction

1.2 Notion

- A rating function has a high discriminative power if it is able to distinguish between debtors with high and with low creditworthiness
- Criteria that serve to evaluate discriminative power:
  - Spread of probability forecasts for the future situation (default or non-default)
  - Separation of the two groups of debtors remaining solvent and those becoming insolvent
  - Concentration of groups of debtors (remaining solvent or becoming insolvent) in good or bad rating classes

Agenda

2 Measures of Discriminative Power

2.1 Contingency Table

- Frequencies of combinations of occurrences of characteristics (2 rating classes)

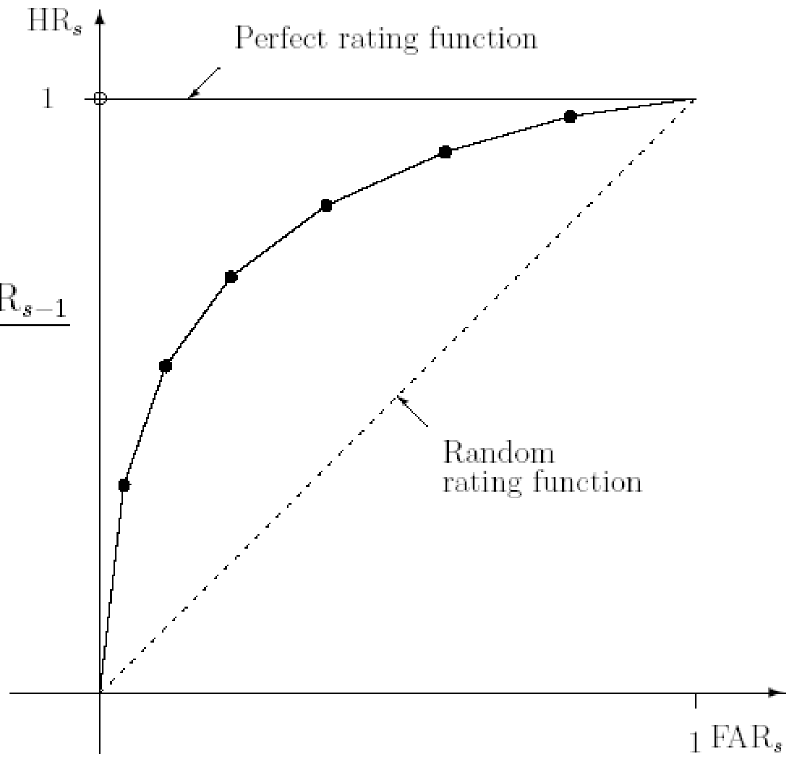
		Observation in $t = 1$	
		Default	Non-default
Prognosis in $t = 0$	Default	A	B
	Non-default	C	D

- Hit rate: A/(A+C)
- False alarm rate: B/(B+D)
- The contingency table only allows an isolated assessment of the ability to separate with regard to the cut-off point chosen

2.2 ROC and AUC

- Receiver operating characteristic (ROC): plot the hit rates (HR) against the false alarm rates (FAR) for all possible cut-off points
- Criterion:
  area under curve (AUC)
- Calculation (discrete):

$$\begin{aligned} \text{AUC} &= \sum_{s=1}^{n} (\text{FAR}_s - \text{FAR}_{s-1}) \cdot \frac{\text{HR}_s + \text{HR}_{s-1}}{2} \\ \text{where} \quad \text{FAR}_0 &= \text{HR}_0 \equiv 0 \\ \text{and} \quad \text{FAR}_n &= \text{HR}_n \equiv 1. \end{aligned}$$



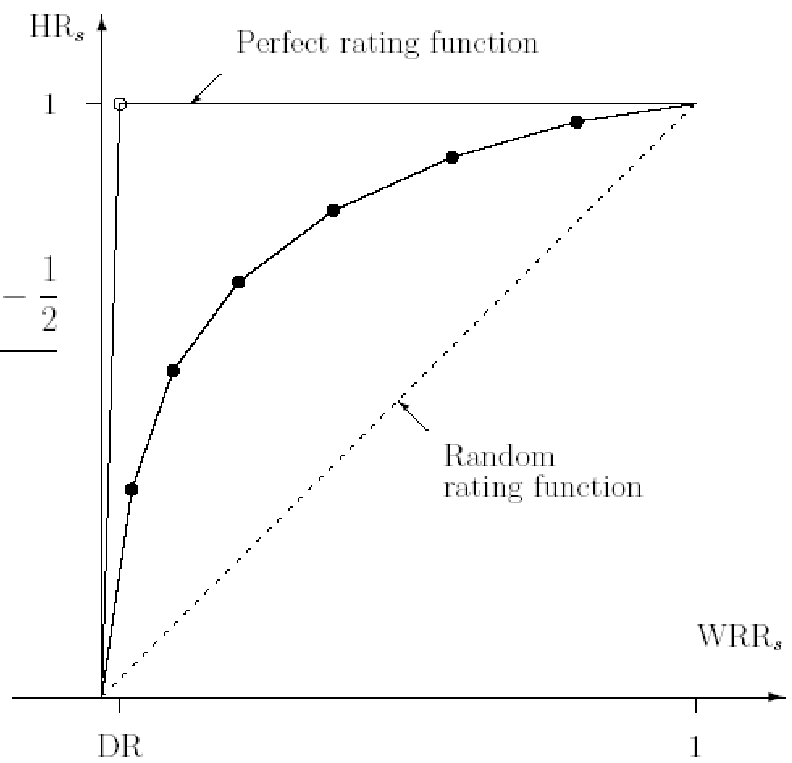
- Identical default rates of compared portfolios are essential

See Sobehart / Kennen (2001) and Engelmann / Hayden / Tasche (2003)

2.3 CAP and AR

- Cumulative accuracy profile (CAP): what fraction of debtors rated worst (WRR) exhibits what fraction of insolvencies (HR)
- Criterion:
  accuracy ratio (AR)
- Calculation (discrete):

$$\mathrm{AR} = \frac{\left(\sum_{s=1}^{n} (\mathrm{WRR}_s - \mathrm{WRR}_{s-1}) \cdot \frac{\mathrm{HR}_s + \mathrm{HR}_{s-1}}{2}\right) - \frac{1}{2}}{\frac{1-\mathrm{DR}}{2}}$$



- AR = 2 × AUC - 1
- Identical default rates (DR) of compared portfolios are essential

See Sobehart / Kennen / Stein (2000) and Engelmann / Hayden / Tasche (2003)

2.4 Stochastic Tendency and Relative Effect

- The false alarm rate distribution tends to be stochastically smaller (or bigger) than the hit rate distribution if

$$\underbrace{\sum_{s=1}^{n} \frac{\mathrm{HR}_s + \mathrm{HR}_{s-1}}{2} \cdot \mathrm{far}_s}_{\equiv \mathrm{RE}} \;<\; (>) \;\frac{1}{2} \quad \text{with} \quad \mathrm{far}_s = \mathrm{FAR}_s - \mathrm{FAR}_{s-1}.$$

- The relative effect does not react to a transformation *m* that maintains order

$$\mathrm{RE} = \mathrm{Prob}\Big(m\left(S^{\mathrm{d}}\right) < m\left(S^{\mathrm{nd}}\right)\Big) + \frac{1}{2} \cdot \mathrm{Prob}\Big(m\left(S^{\mathrm{d}}\right) = m\left(S^{\mathrm{nd}}\right)\Big)$$

2.5 Stochastic Dominance

- Used for a comparison of the hit rate and false alarm rate distribution of **1** rating function *(Use A)*
- Used for a comparison of the ROC curves of **2** rating functions interpreted as distributions *(Use B)*

2.5 Stochastic Dominance

1st order stochastic dominance *(Use A)*

- There is 1st order stochastic dominance of the false alarm rate distribution over the hit rate distribution if the following applies for the hit rates and the false alarm rates

$$\mathrm{FAR}_s \leq \mathrm{HR}_s \quad \forall s$$

The inequality needs to be strictly fulfilled for at least one rating class *s.*

2.5 Stochastic Dominance

1st order stochastic dominance *(Use B)*

- The ROC curve of rating function *R* dominates the curve of function *T* in case of identical false alarm rates if the following inequality is strictly fulfilled for at least one rating class *s*

$$\mathrm{HR}_T(s) \leq \mathrm{HR}_R(s) \quad \forall s$$

2.5 Stochastic Dominance

2nd order stochastic dominance *(Use A and B)*

- *Use A*

$$\sum_{s=1}^{i} \mathrm{FAR}_s \leq \sum_{s=1}^{i} \mathrm{HR}_s \quad \forall i$$

- *Use B* (in case of identical false alarm rates)

$$\sum_{s=1}^{i} \mathrm{HR}_T(s) \leq \sum_{s=1}^{i} \mathrm{HR}_R(s) \quad \forall i$$

2.5 Stochastic Dominance

Example *(Use A)*

Score	Issuers	Defaults	hr	far	HR	FAR	$\sum_{i=1}^{s} HR_s$	$\sum_{i=1}^{s} FAR_s$
1	160	100	33 %	20 %	33 %	20 %	0,33	0,20
2	40	30	10 %	3 %	43 %	23 %	0,77	0,43
3	200	30	10 %	57 %	53 %	80 %	1,30	1,23
4	200	140	47 %	20 %	100 %	100 %	2,30	2,23

2nd order stochastic dominance, but AUC = 47 %

Agenda

1. Introduction
2. Measures of Discriminative Power
3. Empirical Findings
4. Summary

3 Empirical Findings

3.1 Questions

- What discriminative power do the ratings of the agencies Standard & Poor's (S&P) and Moody's Investors Service (Moody's) have?
  - ➢ Area under curve

- Is there (ongoing) stochastic dominance of the rating function of one agency?
  - ➢ 2nd order stochastic dominance

3 Empirical Findings

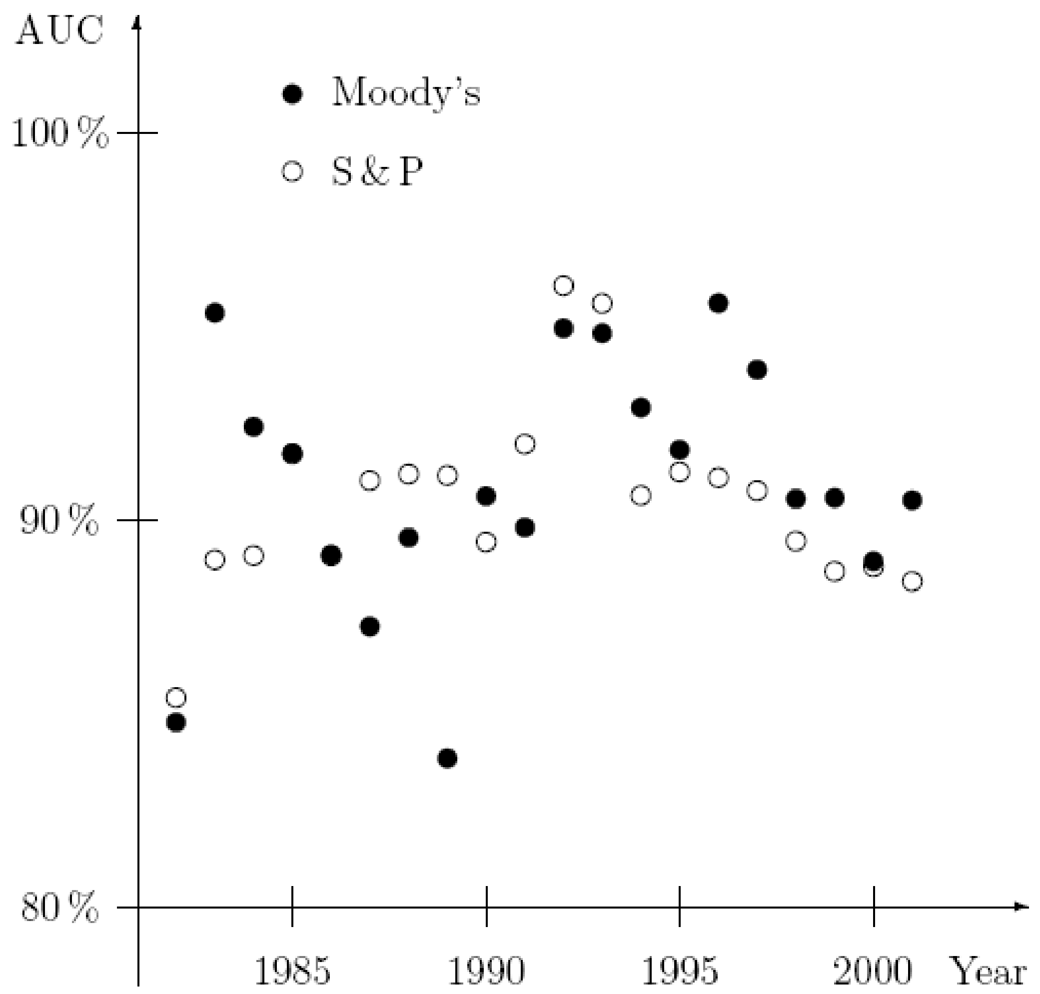
3.2 Data

- Number of issuers and default rates from 1982 to 2001
- For the 7 classes of both, S&P's rating scale (AAA, AA ... CCC) and Moody's (Aaa, Aa ... C)
- **Moody's data:** issuers of long-term bonds from the sectors of industry, transport, providers, and financial institutions (no issuers of structured products, no public issuers)
- **S&P data:** issuers of long-term bonds from the sectors of industry, providers, financial institutions, and insurance (no issuers of structured products, no public issuers)

3 Empirical Findings

3.3 Area Under Curve

- All calculated AUC values are highly significant using Mann-Whitney
- Average area under curve for S&P is 90.5 percent, for Moody's it amounts to 90.9 percent
- No significant difference



3 Empirical Findings

3.4 Stochastic Dominance

- In 17 of the 20 years examined a ranking of the rating functions is possible according to the criterion of 2nd order stochastic dominance
- In 3 of the 20 years examined we observe different rankings of the 2nd order stochastic dominance criterion of the hit rates and the AUC criterion

$$\text{FAR difference} \equiv \sqrt{\frac{1}{6}\sum_{s=1}^{6}\left(\text{FAR}_s^{\text{S\&P}} - \text{FAR}_s^{\text{Moody's}}\right)^2}$$

Year	2^{nd}-order hit rate stochastic dominance	AUC dominance	FAR difference
1982	none	S & P	4.1 %
1983	Moody's	Moody's	1.5 %
1984	none	Moody's	2.7 %
1985	S & P	Moody's	2.1 %
1986	S & P	Moody's	2.9 %
1987	S & P	S & P	3.6 %
1988	S & P	S & P	3.3 %
1989	S & P	S & P	2.8 %
1990	S & P	Moody's	2.2 %
1991	S & P	S & P	1.8 %
1992	S & P	S & P	2.4 %
1993	S & P	S & P	2.4 %
1994	none	Moody's	2.5 %
1995	Moody's	Moody's	2.6 %
1996	Moody's	Moody's	2.4 %
1997	Moody's	Moody's	2.6 %
1998	Moody's	Moody's	2.1 %
1999	Moody's	Moody's	1.7 %
2000	Moody's	Moody's	2.7 %
2001	Moody's	Moody's	3.4 %

Agenda

1 **Introduction**

2 **Measures of Discriminative Power**

3 **Empirical Findings**

4 **Summary**

4 Summary

- Measures of concentration or discriminative power can be defined according to the Gini coefficient. When assessing rating functions the accuracy ratio and the area under curve are used
- These measures can be transferred into each other
- Since distributions are compared the criteria of stochastic dominance can be applied
- Empirical results:
  - ➢ The rating functions of S&P and Moody's show high discriminative power for the years 1982 to 2001
  - ➢ An ongoing dominance of one rating agency could not be ascertained