

Простая нечеткая модель кредитоспособности российских публичных компаний

Ивлиев Сергей,

Руководитель направления
решений для финансовых
институтов
компании «Прогноз»

доцент кафедры инструментальных
средств и математических методов в
экономике Пермского
государственного университета, к.э.н.

Предпосылки создания модели...

За 2009-й год просрочка утроилась...

При фактически неизменной в 2009 году величине совокупного кредитного портфеля нефинансовым организациям (**12,5 трлн. руб.**) наблюдается многократный рост просроченной задолженности до уровня **6%** от портфеля*.

Средний уровень резервов вырос с **4,5%** до **8,6%**, при этом доля резерва по безнадежным ссудам в общем объеме резервов выросла с **33,8%** до **52,2%***.

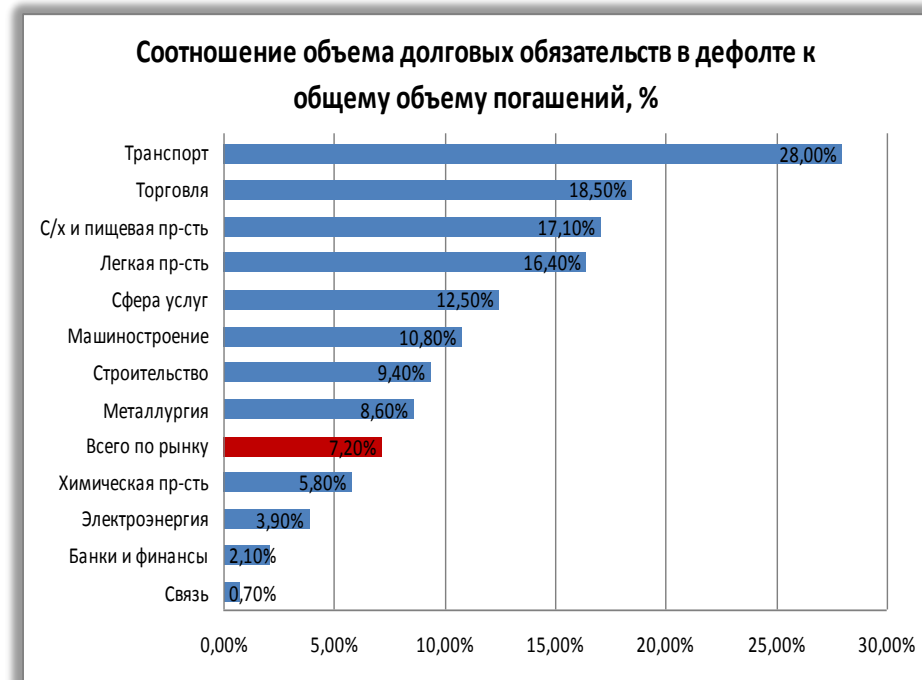
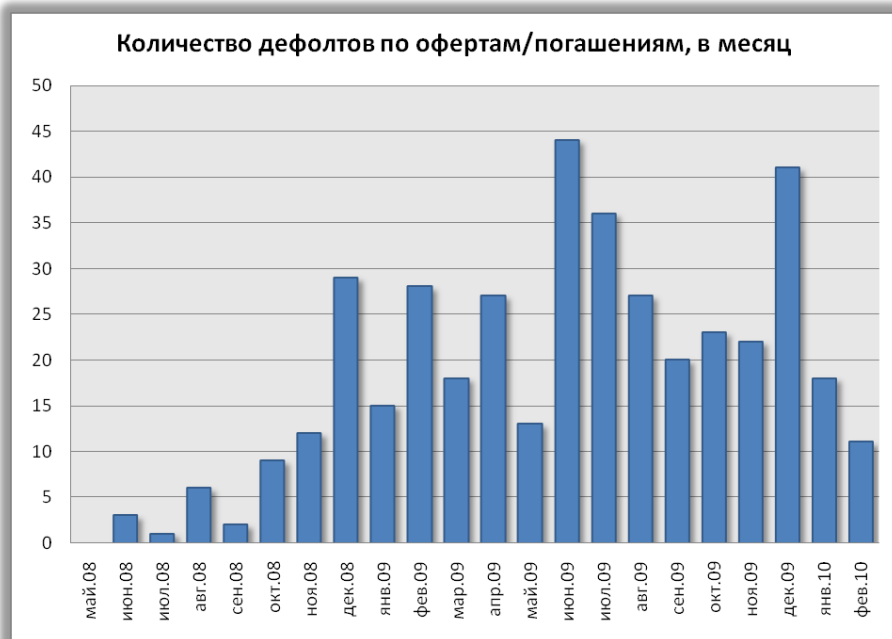


* По данным Обзора банковского сектора РФ №87 январь 2010 года, ЦБ РФ.

Предпосылки создания модели...

2009 год стал рекордным и по дефолтам на рынке облигаций...

До 40 событий в месяц. Около 25% эмитентов рублевых облигаций в дефолте.



* По данным ИБ «Траст» (<http://trust.ru/investment/analitika>).

Требования к модели

...и это хорошо!

(с точки зрения обучающей выборки для построения моделей)

Требования к модели:

1. Модель должна описывать дефолты, которые произошли на рынке облигаций
2. Модель должна быть основана на квартальной финансовой отчетности по РСБУ, чтобы её можно было бы использовать для оценки кредитоспособности в том числе и закрытых компаний
3. Модель должна быть сопоставлена со шкалой международных рейтингов (mapping)
4. Модель должна быть относительно простой, чтобы её можно было широко применять

Выборка

Выборка состоит из данных финансовой отчетности 126 российских компаний-эмитентов рублевых облигаций (*36% от общего количества эмитентов*).

Выборка дефолтов включает 25 событий дефолта по облигациям в период 2008-2009 гг. (*30% от общего количества дефолтов*).

36 компаний имеет присвоенный кредитный рейтинг (*29% от общего количества, для сравнения в генеральной совокупности - 34%*).

Компании-SPV в выборку не включались.

Источник данных: Сервисы раскрытия информации

<http://emitent.prime-tass.ru/EmitentPages/EmitentSearch.aspx>

<http://www.e-disclosure.ru/>

<http://disclosure.fcsм.ru/>

“Bad” Cases & Cure Events

Наблюдение (отчетная дата) относится к классу «плохих» (Def=1), если компания в первый раз объявила дефолт по своему долгу после выхода отчетности.

В соответствии с российским законодательством используются следующие лаги: 30 дней для квартальных отчетов, 90 дней для годовых.

Рассматриваются только реальные дефолты. Технические дефолты (Cure Events) не учитываются.

Источник данных: TRUST Interactive: Defaults Review

<http://trust.ru/investment/analitika/interactive/defaults/rus/index.html>

Объем выборки и период наблюдения

Период наблюдения: 1Q 2008 – 3Q 2009.

Объем выборки: 588 точек (компания * квартал)
(в среднем 4.6 точек на компанию).

В т.ч.:

объем выборки (“Good”): 563 точки

объем выборки (“Bad”): 25 точек (4.25%)

Период прогнозирования: 0-60 дней.

Метод обработки пропусков: case-wise

Обработка данных отчетности

Отчетность (ф.1 «Баланс», ф.2 «Отчет о прибылях и убытках») загружалась в ПК «ПРОГНОЗ. Кредитный риск» и на её основе рассчитывались финансовые коэффициенты, характеризующие кредитоспособность:

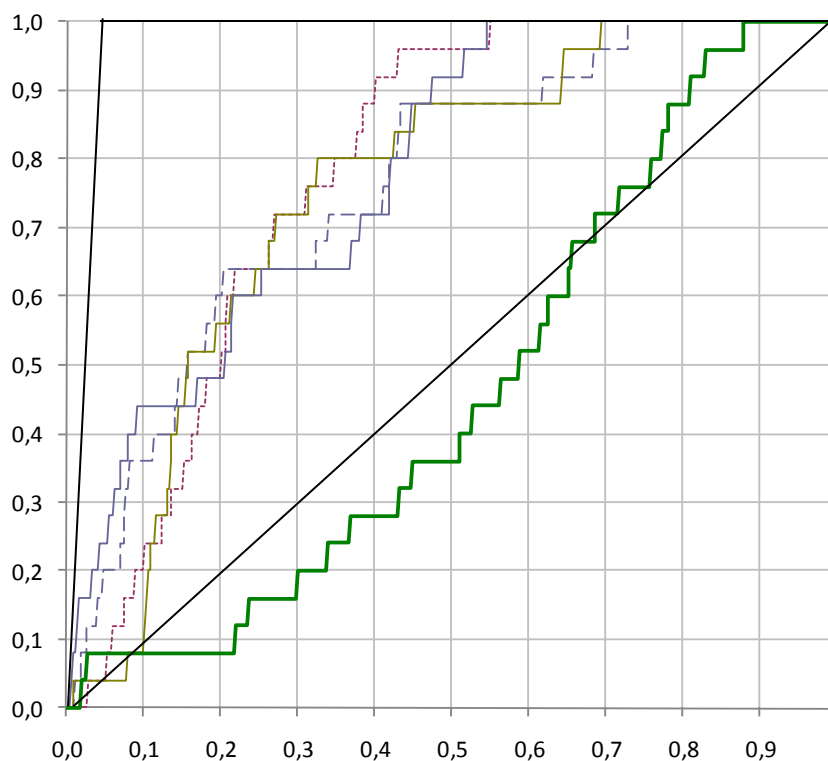
Size	Balance sheet structure	Profitability	Liquidity
LN (Assets)	Working Capital / Assets	EBIT / Sales	Cash / ST Debt
LN (Sales)	Retained Earnings / Assets	EBIT / Assets	Cash and equivalents/ ST Debt
Sales/Assets	Equity / Total Liabilities	EBIT / Interest	Current Assets/ ST Debt

...а также несколько известных моделей (Альтман, Фулмер и др.)

Анализ предсказательной силы переменных

Для каждого из коэффициентов была построена ROC-кривая и рассчитан коэффициент Джини (in-sample Gini accuracy ratio).

Лучшими предикторами оказались:



- - - EBIT / I
 - - - LN(Sales)
 — Retained Earnings / Assets — Equity / Total Liabilities
 — Altman Z''-Score

Variable	In-sample Gini AR
Equity / Total Liabilities	57.8%
EBIT / Interest	55.8%
LN(Sales)	54.3%
Retained Earnings / Assets	52.3%

По остальным оценки хуже, например:

Variable	In-sample Gini AR
EBIT / Assets	39.6%
Sales / Assets	23.3%
EBIT / Sales	10.6%
Altman's Z''-Score (EM Score)	-7.7%

Корреляционный анализ

Для исключения мультиколлинеарности были рассчитаны парные коэффициенты линейной корреляции и рассчитаны собственные значения матрицы корреляций (МГК).

	LN (Sales)	Ret Earnings / Assets	Equity / Liabilities
EBIT / I	0,1160	0,3015	0,3726
LN (Sales)		0,2104	0,3507
Ret Earnings / Assets			0,133

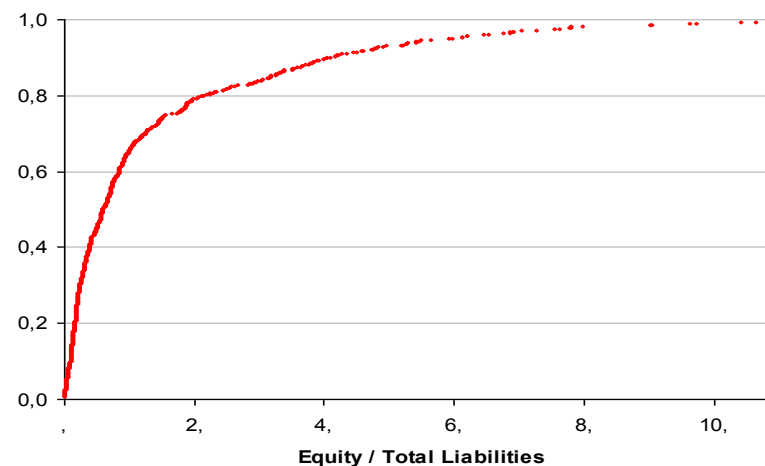
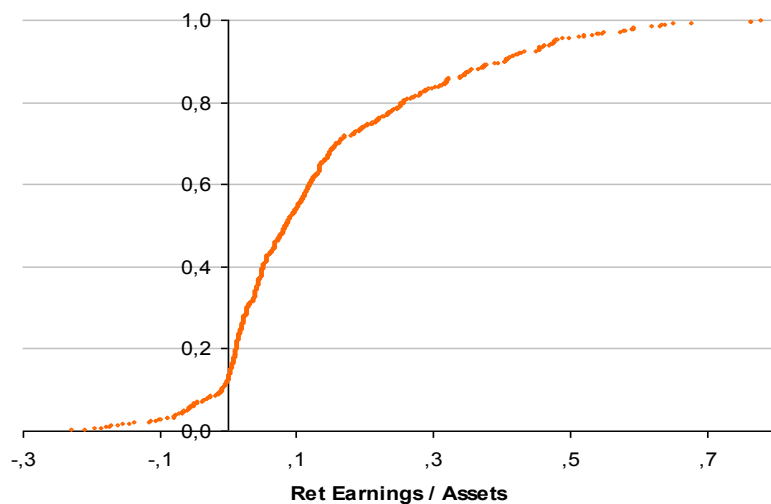
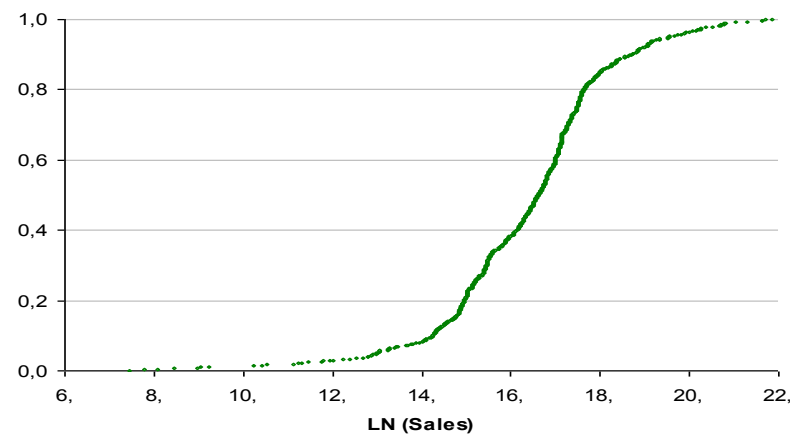
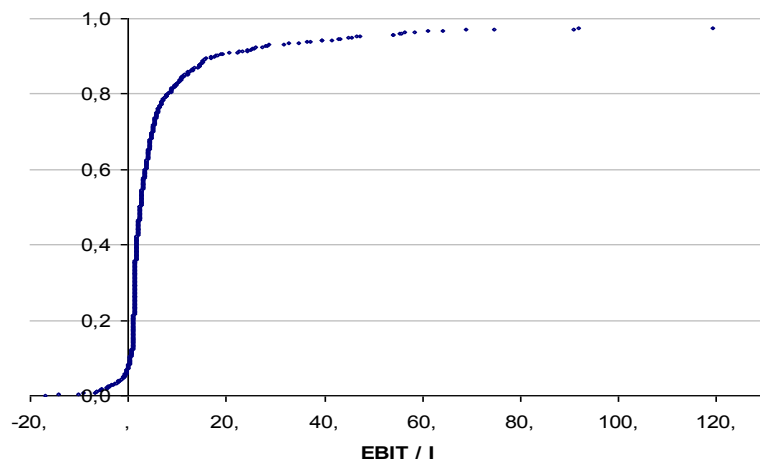
Коэффициенты корреляции не превышают 0,4.

Первое собственное значение не превышает 50% (доля объясняемой первой главной компонентой вариации).

Таким образом, не имеет смысла исключать переменные или переходить к главным компонентам.

Анализ распределения

Для анализа вероятностных свойств переменных были построены эмпирические функции распределения:



Требования к модели

Требования к модели:

1. Модель должна быть простой, чтобы её можно было легко объяснить и использовать в практике
2. Предсказывающая сила модели должна быть выше, чем у отдельных переменных
3. Модель должна корректно учитывать отсутствие нормальности в исходных переменных

Рассматриваемые варианты:

1. Простая модель с порогами (S-Score)
2. Простая модель на основе нечетких множеств (FS-Score*)
3. Логит-модель на основе исходных переменных (Logit)
4. Логит-модель на основе преобразованных переменных (Logit F)

Простая модель с порогоми (S-Score)

Спецификация модели S-Score:

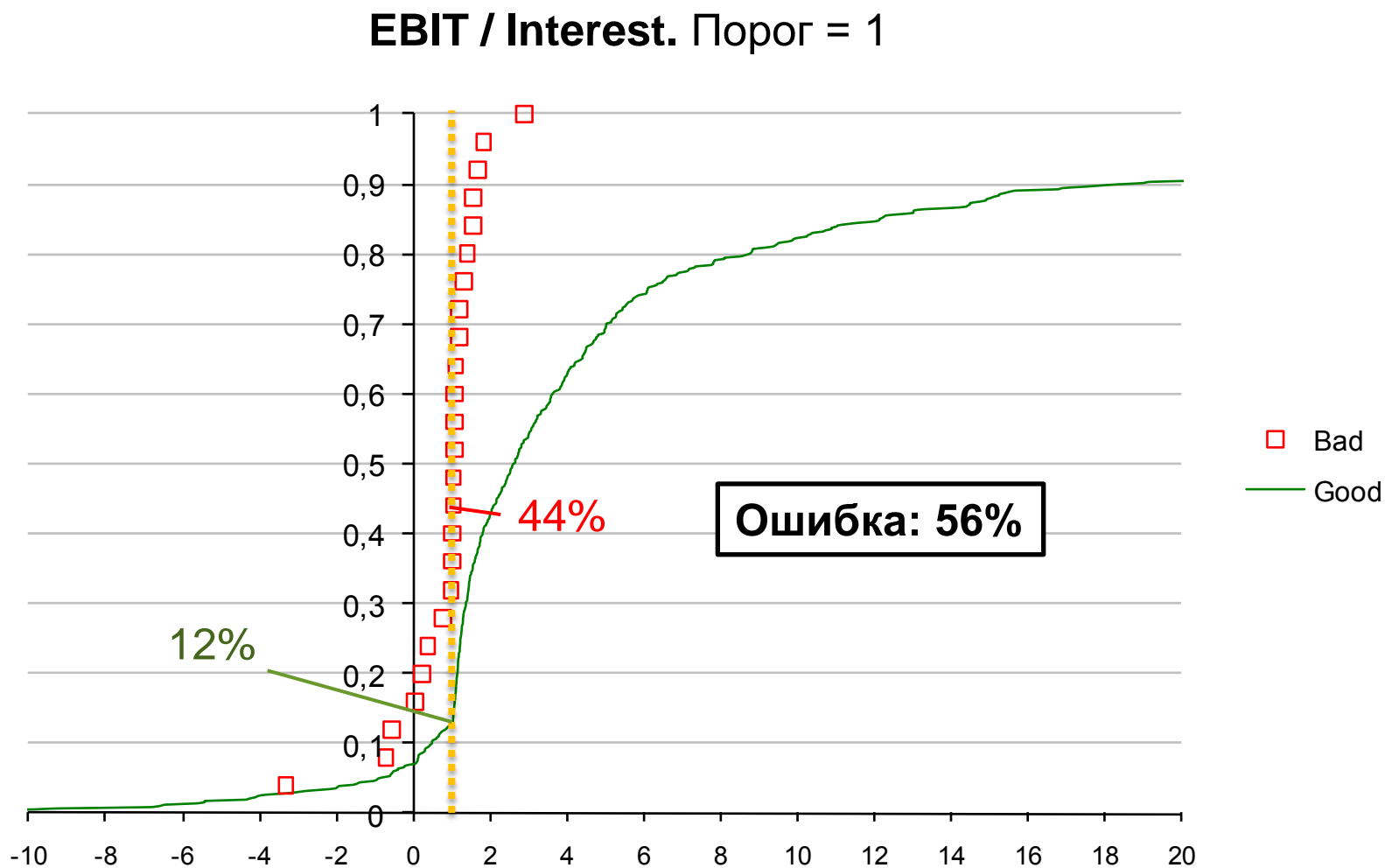
$$S = \sum_{i=1}^n 1_{\{X_i > c_i\}}$$

где X_i – исходная переменная, c_i – порог отсечения.

Пороги отсечения калибровались, чтобы обеспечить минимальную суммарную ошибку классификации (ош. I рода + ош. II рода).

Простая модель с порогоми (S-Score)

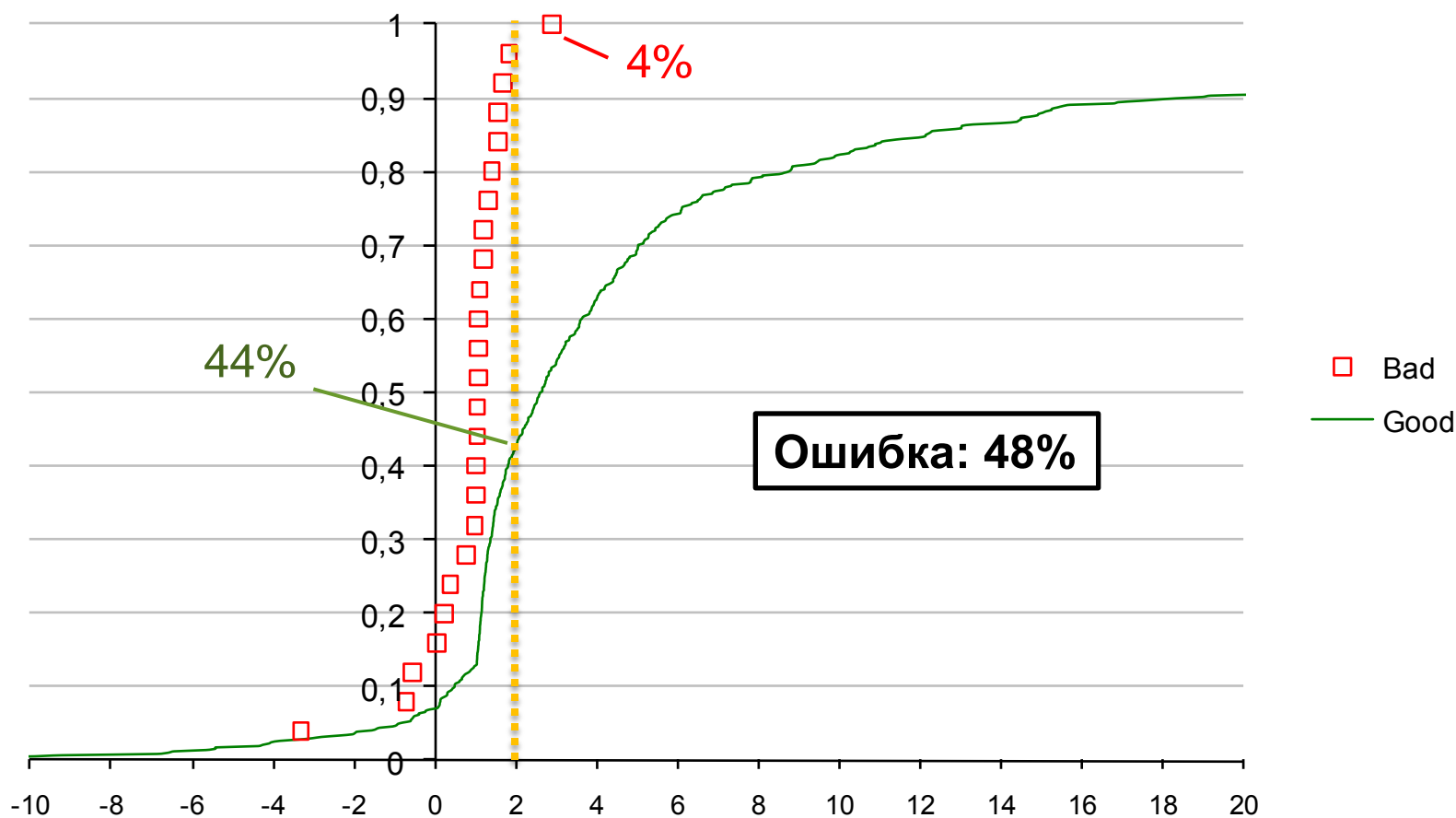
Иллюстрация на примере...



Простая модель с порогоми (S-Score)

Иллюстрация на примере...

EBIT / Interest. Порог = 2



Простая модель с порогоми (S-Score)

Результаты:

Оптимальные пороги отсечения и ошибка классификации :

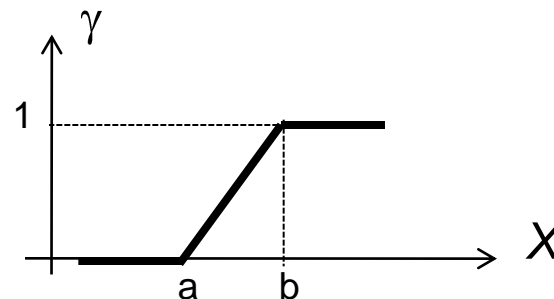
Variable	Best cut-off	Type I error	Type II error	Total error
EBIT / Interest	2	44%	4%	48%
LN(Sales)	16	37%	28%	65%
Retained Earnings / Assets	0.04	33%	20%	53%
Equity / Total Liabilities	0.5	44%	12%	56%
S-Score	1	27%	8%	35%

Предсказательная сила модели (in-sample Gini AR): **71,8%**

Модель на основе нечетких множеств (FS-Score)

Спецификация модели FS-Score (Fuzzy Simple Score):

$$\gamma_i(X_i, a_i, b_i) = \begin{cases} 0, & X_i < a_i \\ \frac{X_i - a_i}{b_i - a_i}, & a_i \leq X_i < b_i \\ 1, & X_i \geq b_i \end{cases}$$



$$FS = \sum_{i=1}^n \gamma_i(X_i, a_i, b_i)$$

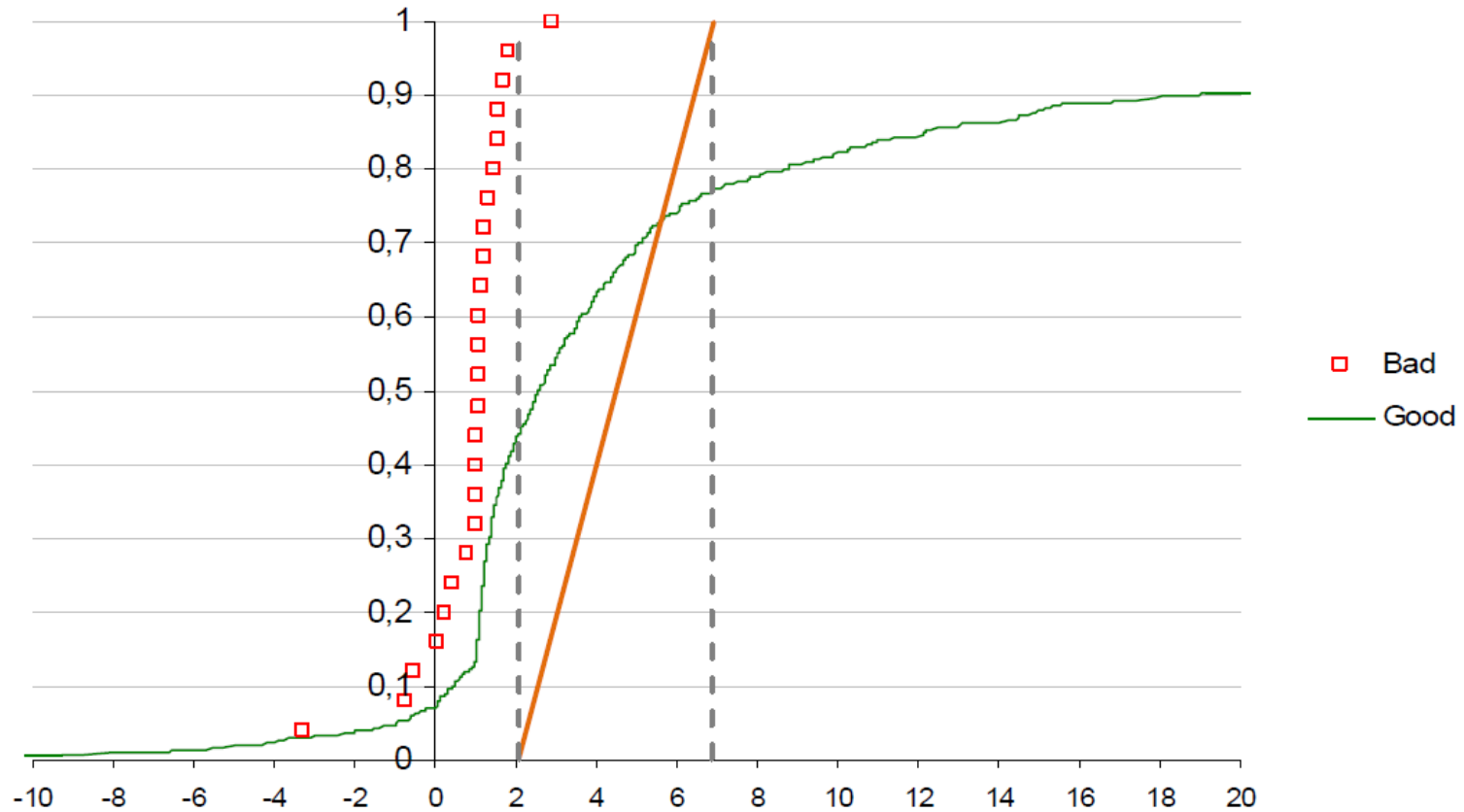
где γ_i - функция принадлежности к множеству «хороших»,
 X_i – исходная переменная, a_i, b_i – пороги.

a_i можно взять из предыдущей модели, b_i выбираем из экономических соображений, так, чтобы $\gamma=1$ соответствовала наиболее кредитоспособным компаниям.

Модель на основе нечетких множеств (FS-Score)

Иллюстрация на примере...

EBIT / Interest. Пороги $a = 2$, $b = 7$



Модель на основе нечетких множеств (FS-Score)

Результаты:

Выбранные пороги:

Variable	a_i cut-off	b_i cut-off
EBIT / Interest	2	7
LN(Sales)	16	18
Retained Earnings / Assets	0.04	0.2
Equity / Total Liabilities	0.5	2

Предсказательная сила модели (in-sample Gini AR): **72,7%**

Логит-модель

Спецификация модели Logit:

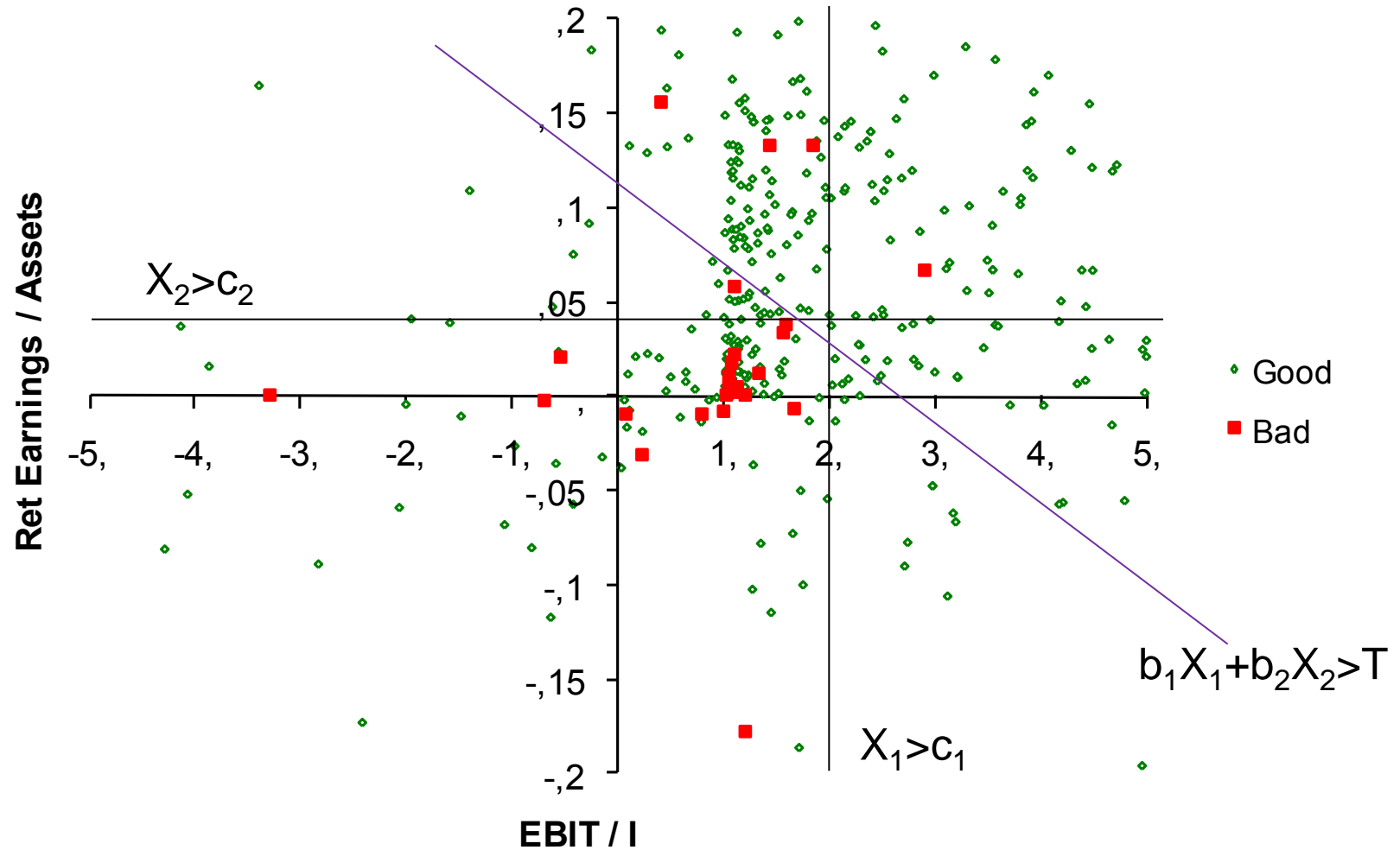
$$y = \frac{e^{b_0 + \sum_{i=1}^n b_i X_i}}{1 + e^{b_0 + \sum_{i=1}^n b_i X_i}}$$

где y - бинарная переменная (0 для «хороших», 1 для «плохих»),
 X_i – исходная переменная, b_i – коэффициенты регрессии.

Оценивание: методом максимального правдоподобия.

Логит-модель

Иллюстрация на примере...



Логит-модель

Результаты:

Коэффициенты регрессии:

Variable	i	b_i
Const	0	1,9808
EBIT / Interest	1	-0,1131
LN(Sales)	2	-0,2431
Retained Earnings / Assets	3	-3,1491
Equity / Total Liabilities	4	-2,0711

Предсказательная сила модели (in-sample Gini AR): **70,5%**

Логит-модель на нечетких переменных

Спецификация модели и результаты Logit F:

$$y = \frac{e^{b_0 + \sum_{i=1}^n b_i \gamma_i}}{1 + e^{b_0 + \sum_{i=1}^n b_i \gamma_i}}$$

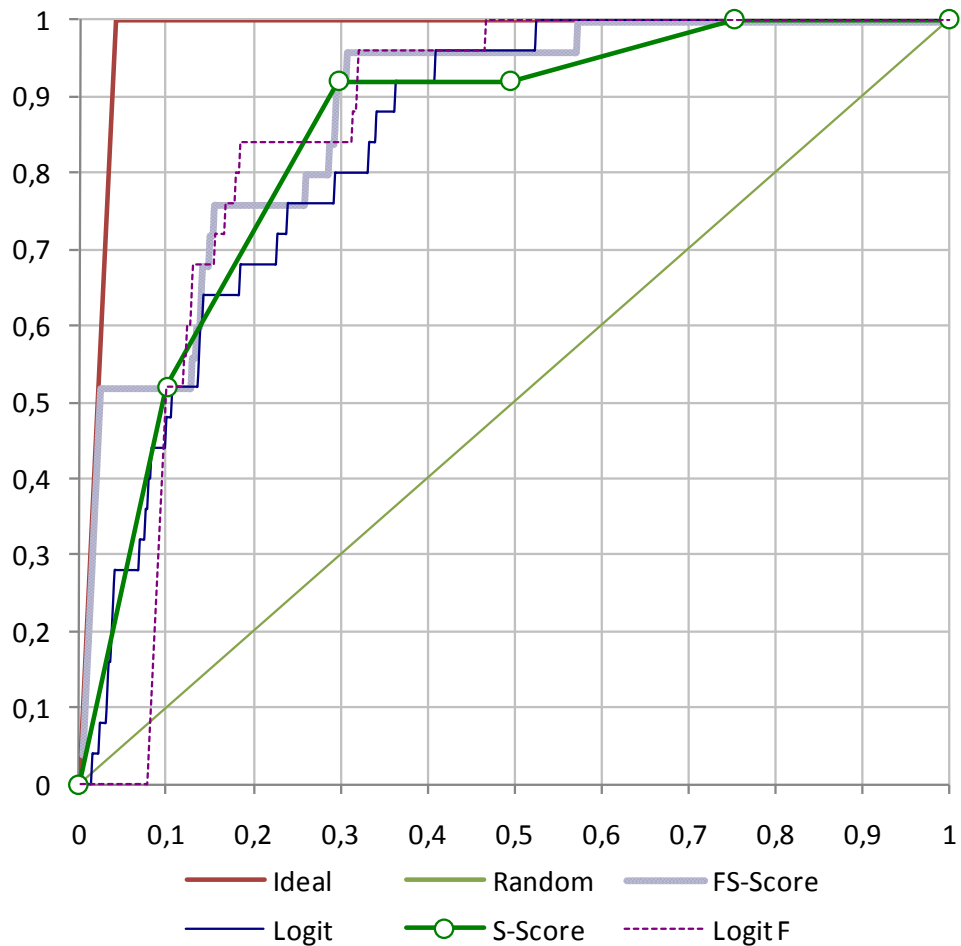
где y - бинарная переменная (0 для «хороших», 1 для «плохих»),
 γ_i – нечеткое число для X_i , b_i – коэффициенты регрессии.

Variable	i	b_i
Const	0	-1,46645
EBIT / Interest	1	-6,21185
LN(Sales)	2	-1,19298
Retained Earnings / Assets	3	-3,1798
Equity / Total Liabilities	4	-5,09643

Предсказательная сила модели (in-sample Gini AR): **72,9%**

Сравнение моделей

Модели в целом близки по предсказательной силе... FS-Score и Logit F выглядят предпочтительнее. FS-Score при этом ещё и проще...



Маппинг с внешними рейтингами

Для практического применения с целью оценки вмененного рейтинга, а также с целью валидации, был произведен маппинг к присвоенным внешним рейтингам (S&P, Moodys', Fitch Ratings)...

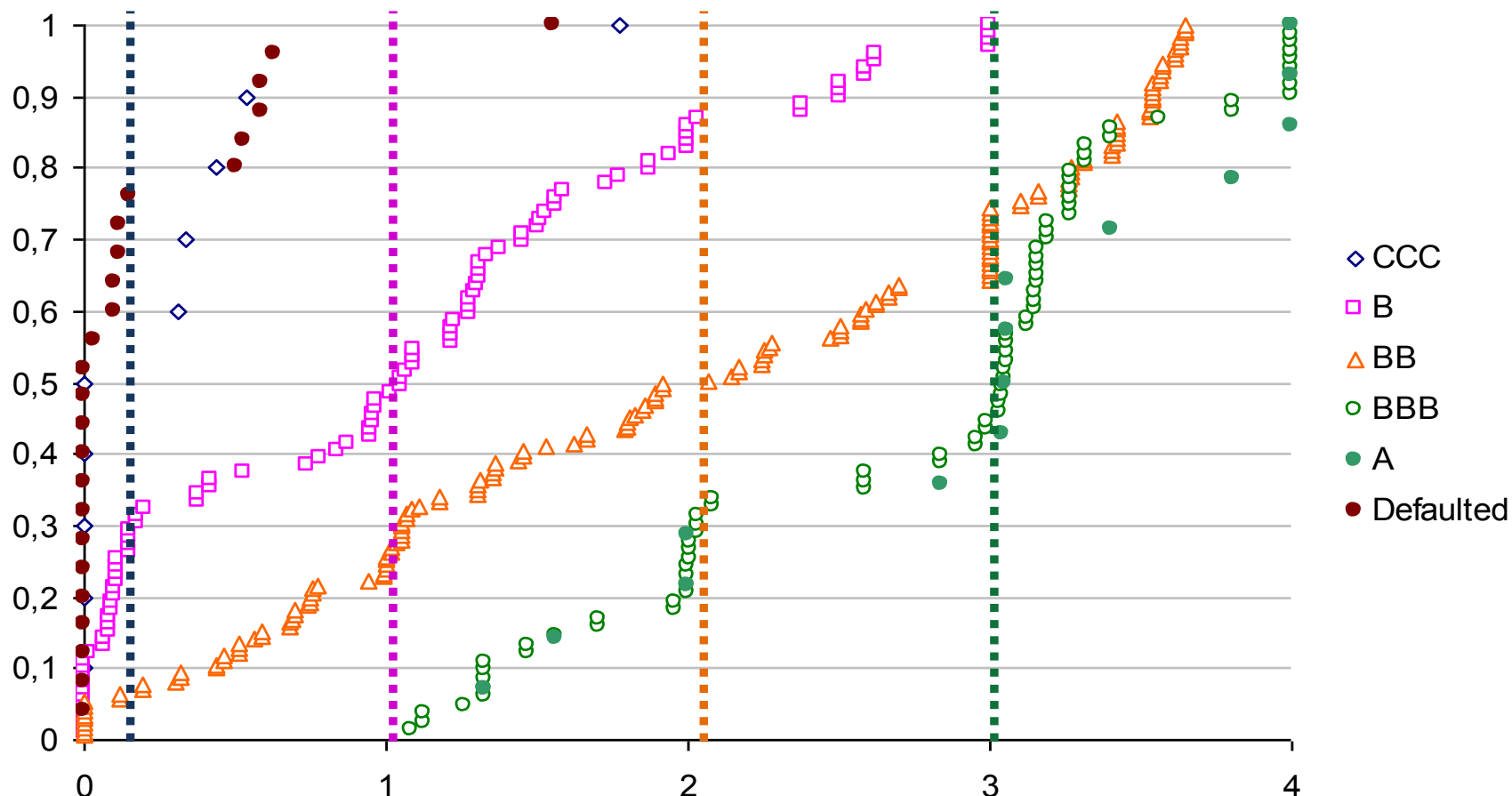
Шкала была редуцирована до 5 классов: A, BBB, BB, B, CCC/C.

Рассчитывались значения FS-Score для компаний соответствующего класса рейтинга.

Rating grade	25% Quintile	Median	75% Quintile
A	2	3.06	3.81
BBB	2.01	3.05	3.27
BB	1	2.11	3.1
B	0.11	1.05	1.56
CCC/C	0	0.15	0.44
Defaulted	0	0	0.15

Маппинг с внешними рейтингами

Эмпирические функции распределения FS-Score для различных классов рейтинга... Медианы классов значительно различимы...



Эмпирическое правило: FS-Score равно количеству букв В в рейтинге

Результаты и перспективы развития

Результаты:

Полученная модель *FS-Score* с одной стороны позволяет достаточно адекватно описать произошедшие в 2008-2009 гг дефолты на рынке облигаций (*in-sample Gini AR = 72,7%*), с другой стороны является относительно простой, чтобы её можно было применять в широкой практике.

Дальнейшие направления развития модели:

- ✓ Расширение набора предикторов
- ✓ OOS валидация на свежих дефолтах Q4 2009-Q1 2010
- ✓ Анализ устойчивости модели во времени и в разрезе секторов
- ✓ Применение множественной логит-регрессии с целью повысить качество маппинга внешних рейтингов
- ✓ Учет весов ошибок I и II рода при калибровке пороговых значений
- ✓ Применение копулы для более эффективного учета совместного распределения предикторов

Спасибо за внимание!

Препринт:

Simple Fuzzy Score for Russian Public Companies Risk of Default.

Preprint published at Cornell University Library Arxiv.

<http://arXiv.org/abs/1004.0685>

Контактная информация:

email ivliev@prognoz.ru



sivliev