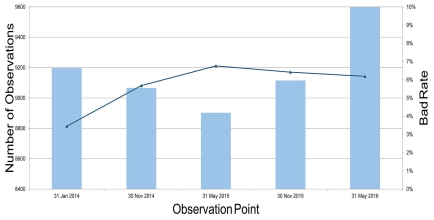
Probability of Default for SME Loans

by K. Angelova, Y. Tsoneva and E. Uzunova

Total observations - 45 881
Explanatory variables - 132 - but no data for obligator risk characteristics
Types of products - Credit loans and cards, Overdrafts, etc.

Good/Bad - standard definition: the obligor is past due more than 90 days on any material credit obligation within 12 months.

Number of observations and Bad Rate By Observation Sample



Exclusions:
31 Jan 2014

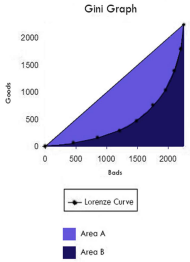
Model Development Sample (70%)					
Bads	Frequency	Percent	Cumulative Frequency	Cumulative %	% of Total cases
0	17767	93.72	17767	93.72	38.72%
1	1191	6.28	18958	100	2.60%
Model Validation Out - of – Sample (30%)					
Bads	Frequency	Percent	Cumulative Frequency	Cumulative %	% of Total cases
0	7615	93.71	7615	93.71	16.60%
1	511	6.29	8126	100	1.11%
Model Validation Out - of – Time (Observation Point 31.05.2016)					
Bads	Frequency	Percent	Cumulative Frequency	Cumulative %	% of Total cases
0	9005	93.81	9005	93.81	19.63%
1	594	6.19	9599	100	1.29%
Exclusions	9198				20.05%

3 samples:
- Development
- Validation Out of Sample
- Validation Out of Time

With Exclusions

Variable	Gini
Overdue_amount	0.58
Number_transactions	0.53
TOTAL_OVDUE_PRINC_1	0.53
Number_trans_3M	0.51
TOTAL_OVDUE_PRINC_2	0.46
TOTAL_OVDUE_PRINC_4	0.44
Total_balance	0.43
Outst_appr	0.43
TOTAL_OVDUE_PRINC_3	0.43
Ovdue_amount_CC_OVER_6M	0.41
TOTAL_OVDUE_PRINC_5	0.40
WORST_GROUP_1	0.39
B_OVDUE_INTEREST_1	0.39
TOTAL_OVDUE_PRINC_6	0.36
B_OVDUE_INTEREST_2	0.35

Good correlations and high Gini coeeficient



Balance Increase 3m	class	goods	% goods	bads	% bads	total	% total	GB Odds	Bad Rate	Distribution Bad (DB)	Distribution Good (DG)	WOE	IV
	1	9044	38%	233	16%	9277	37%	38.82	0.03	0.16	0.38	0.83	
	2	4771	20%	250	18%	5021	20%	19.08	0.05	0.18	0.20	0.12	
	3	5836	24%	359	25%	6195	24%	16.26	0.06	0.25	0.24	-0.04	0.36
	4	4333	18%	571	40%	4904	19%	7.59	0.12	0.40	0.18	-0.81	

** IV above 0.3 suggests strong predictive power*

Analysis of Maximum Likelihood Estimates

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	-2.69	0.04	4835.39	<.0001
avgutil_d_w	0.37	0.06	33.20	<.0001
balance_b_incr_3m	0.42	0.06	47.38	<.0001
Number_transactions	0.54	0.04	189.46	<.0001
Overdue_amount	0.53	0.03	277.84	<.0001
Total_balance	0.28	0.05	30.53	<.0001
WORST_GROUP_1	0.39	0.03	138.38	<.0001
Industry	0.77	0.16	22.98	<.0001

Hosmer and Lemeshow Goodness-of-Fit Test			
			Pr > ChiS
Chi-Square	DF	q	
33.95	8.00	<.0001	

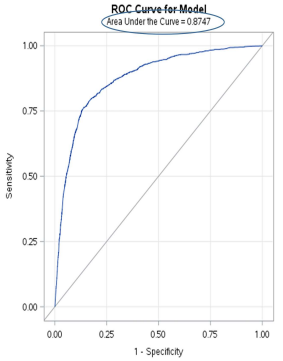
Somers' D	0.75
Gamma	0.75
Tau-a	0.09

Results seemed good with indications for potential non-stationarity or mispecification error.

Score Class	Min Score	Max Score	Median Score	Num cases	Bad	Mean PD	Good	GB odds	ln(GB odds)
1	-68.24	70.11	11.17	947	397	47%	550	1.39	0.33
2	70.75	138.40	104.78	949	262	25%	687	2.62	0.96
3	138.68	222.36	187.62	987	191	14%	796	4.17	1.43
4	222.52	266.56	251.89	914	82	8%	832	10.15	2.32
5	266.57	292.06	276.94	946	43	6%	903	21.00	3.04
6	292.07	329.46	308.80	1506	62	4%	1444	23.29	3.15
7	330.29	379.37	354.28	2555	68	3%	2487	36.57	3.60
8	380.35	429.79	404.24	2772	42	2%	2730	65.00	4.17
9	430.25	479.41	457.09	3834	29	1%	3805	131.21	4.88
10	481.27	528.32	503.96	2528	13	1%	2515	193.46	5.27
11	530.16	576.85	542.56	1020	2	0%	1018	509.00	6.23

Confusion Matrix of Development Sample

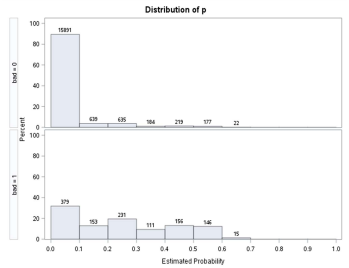
Cut -off point	Correct		Incorrect		Percentages				
Level	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False Positive	False Negative
0.5	159	17568	199	1032	93.5	13.4	98.9	55.6	5.5
0.4	314	17337	430	877	93.1	26.4	97.6	57.8	4.8
0.3	428	17164	603	763	92.8	35.9	96.6	58.5	4.3



Sample	Gini Coefficient	Area under ROC curve
Development	0.749	0.875
Validation out of sample	0.732	0.870
Validation out of time	0.724	0.862

Validation Sample (Out of Sample)
KS 0.16 D 0.35

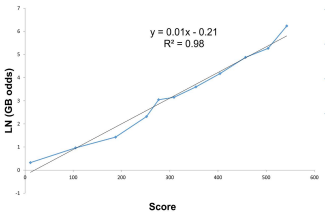
Validation Sample (Out of Time)
KS 0.14 D 0.60



The validations shows some flaus in the model - The KS was very low, however, Gini and ROC curve were showing a good overall fit. The comparative histograms also indicated overlaps, meaning problems with the model.

Finally, we calibrated to S&P and received the results on the right.

The last graph shows that we have a similar trend to the one figurating in the master scale.



Rating	Score	PD
BBB	730	0.4%
BBB-	722	0.6%
BB+	712	1.1%
BB	701	1.7%
BB-	690	2.9%
B+	680	4.4%
B	674	5.9%
B	668	7.7%
B-	655	13.5%
CCC/C	637	25.4%
CCC/C	618	46.7%