

Schätzung von Verlustquoten – Methodische Probleme und Lösungsansätze

Thomas Hartmann-Wendels
Hans-Christian Elbracht
Forschungsinstitut für Leasing an der
Universität zu Köln

Relevanz des Loss Given Defaults

- Verlustquote bei Ausfall (LGD) zentraler Parameter zur Quantifizierung von Kreditrisiken (PD, EAD, Korrelation),
- Einsatzbereiche
 - regulatorische Kapitalanforderungen (Basel II/III, SolvV),
 - MaRisk (Risikoerkennung, -analyse und -steuerung)
 - risikoadjustierte Preisbildung (EL),
 - Allokation von Risikokapital (UL)
- fortgeschrittener IRB-Ansatz: Entwicklung eigener Modelle zur LGD-Schätzung auf Basis historischer Daten,
- IRB-Basisansatz: LGD =45%, für Leasing häufig zu hoch angesetzt,
- Eigenmittelunterlegung und erwarteter Verlust (EL) linear im LGD

Relevanz des Loss Given Defaults

- market recovery vs. workout- recovery
- empirische Arbeiten basieren überwiegend auf market recovery, z.B. Gupton/Stein (2005), Guha (2003), Altman/Gande/Saunders (2004), Hamilton/Gupton/Berthault (2001), Altman/Kishore (1996)
- empirische Arbeiten zur workout recovery: Hamilton/Carty (1999), Schmit/Stuyck (2002), De Laurentis/Riani (2005), Bellotti/Crook (2009), Grunert/Weber (2009), Bastos (2010), Gürtler/Hibbeln (2011)
 - teilweise geringer Datenumfang, Daten stammen ausschließlich von *einem* KI; keine Angaben zur Verteilung der Verlustquoten, wenige Angaben zu möglichen Einflussgrößen auf den LGD, keine Schätzung von Verlustquoten

Workout-LGD – Allgemeine Definition

- Verlustdefinition (§126 SolvV): ökonomischer Verlustbegriff
- Verlustquote gibt an, welcher Anteil am ausstehenden Forderungsbetrag „verloren“ ist:

$$\text{LGD} = \frac{\text{ökonomischer Verlust}}{\text{EAD}} = \frac{\text{EAD} - (\text{Erlöse} - \text{Kosten})}{\text{EAD}} = 1 - \frac{\text{Erlöse} - \text{Kosten}}{\text{EAD}} = 1 - \text{RR}$$

- EAD:
 - Barwert der noch ausstehenden Mindestleasingzahlungen + Barwert des kalkulierten Restwerts + aufgezinste säumige Leasingraten
- Erlöse:
 - Verwertungserlöse aus Leasingobjekt, eingehende Leasingraten, Erlöse aus sonstigen Sicherheiten, sonstige Zahlungen, jeweils diskontiert auf den Ausfallzeitpunkt
- Kosten:
 - direkt zurechenbare Kosten
 - nicht direkt zurechenbare Kosten

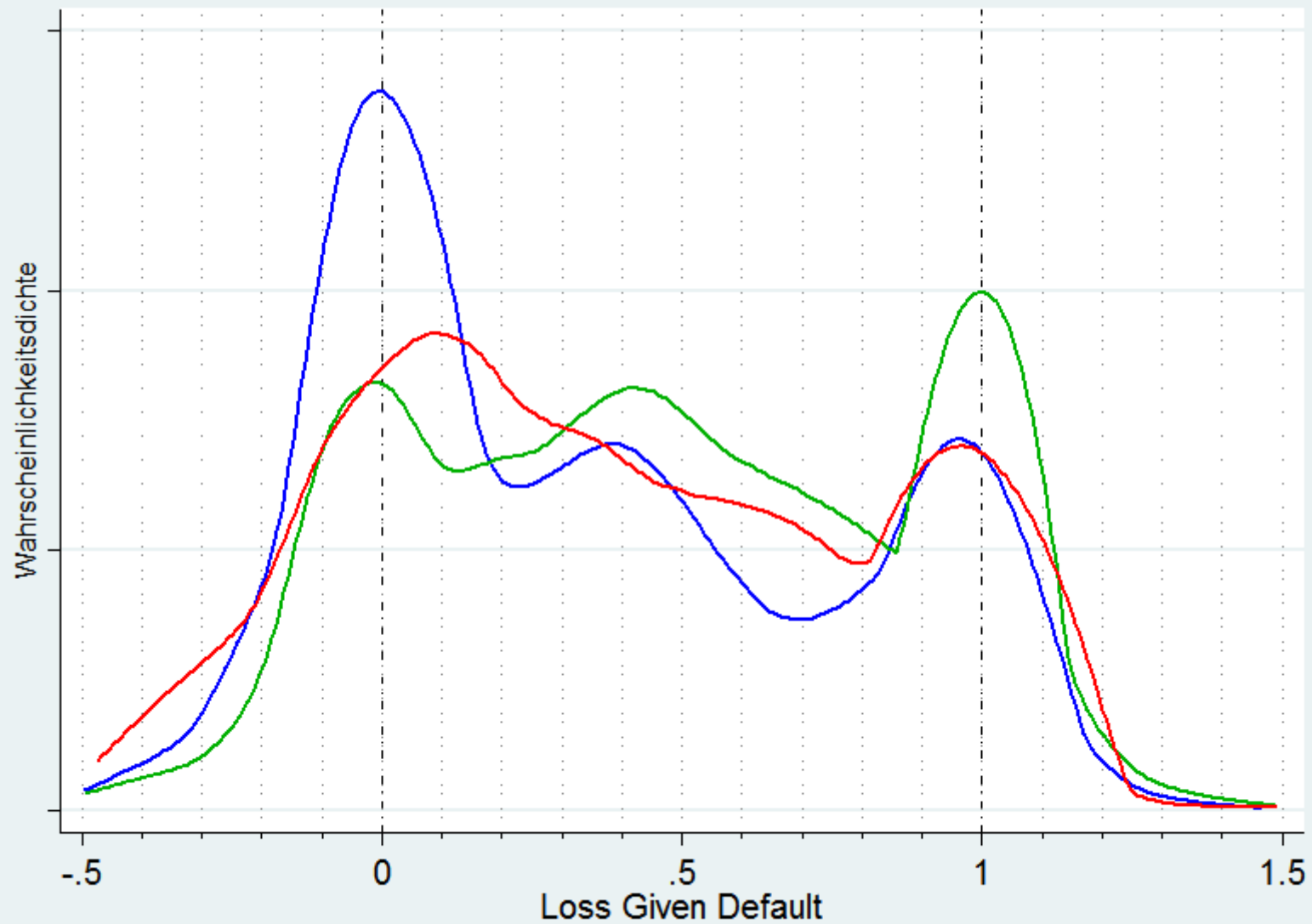
Aspekte, die zu beachten sind

- Ausfalldefinition auf Schuldnerenebene (§125 SolvV)
- Ausfallgründe
 - Insolvenz
 - 90-Tage-Verzug
 - Wertberichtigung
- Wiedergesundung!
- Workout-Prozess → Sanierung oder Abwicklung
- die Vereinfachung LGD aus $[0;1]$ ist nicht immer angemessen
- regulatorische Vorgaben:
 - LGD-Schätzung zu zwei Zeitpunkten,
 - wirtschaftlicher Verlust nicht negativ,
 - Downturn-LGD
 - keine spezifische Vorgaben an Schätzmodelle und Methoden

Datensätze

Zeitraum	2002 – 2010	
Ausgefallene Verträge	27.000	
• davon abgewickelt	25.000	
• davon wiedergesundet	2.000	
Objektkategorien	Anzahl der Verträge	durchschnittlicher LGD
Fahrzeuge	12.400	0.26
Maschinen/Anlagen	10.000	0.35
EDV/Bürotechnik	3.600	0.61
Technische Geräte	500	0.54
Sonstige	250	0.34

Wahrscheinlichkeitsdichte der Verlustquote



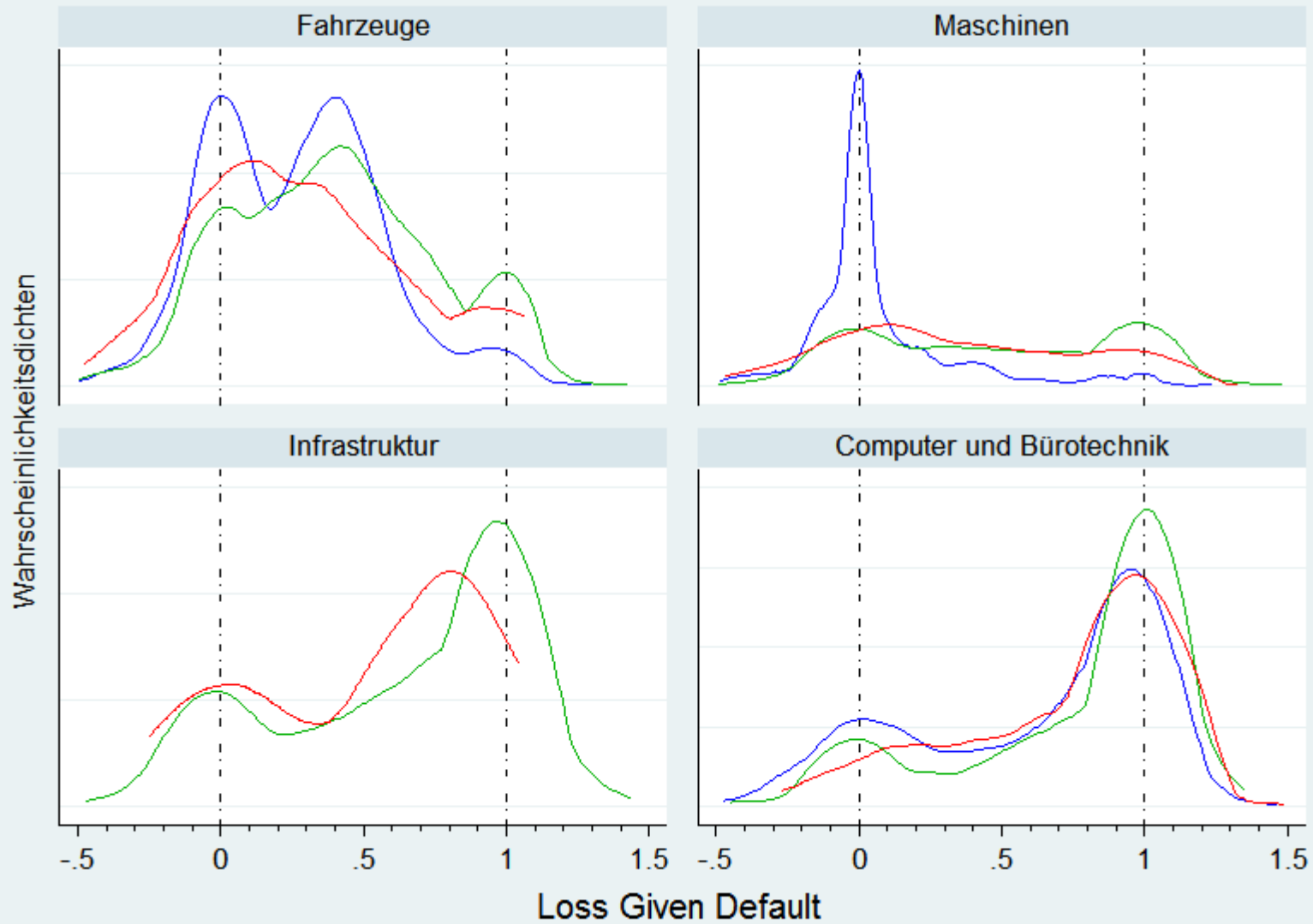
Eigenschaften der Verlustverteilung

- Multi-modale Verteilung der Verlustquoten
→ bleibt (zum Teil) auch erhalten bei Aufspalten nach Objektkategorien
- Streuung der Verlustquoten sehr hoch

Objektkategorie	Interquartilsabstand
Fahrzeuge	0.35 – 0.45
Maschinen/Anlagen	0.45 – 0.85
EDV/Bürotechnik	0.50 – 0.95
Technische Geräte	0.70 – 0.80
Sonstige	0.75 – 0.90

- Verlustquoten < 0 und > 1 sind nicht selten!

Verteilung der Verlustquoten für unterschiedliche Objektkategorien



Verlustquote in Abhängigkeit vom relativen Vertragsalter

Hypothese

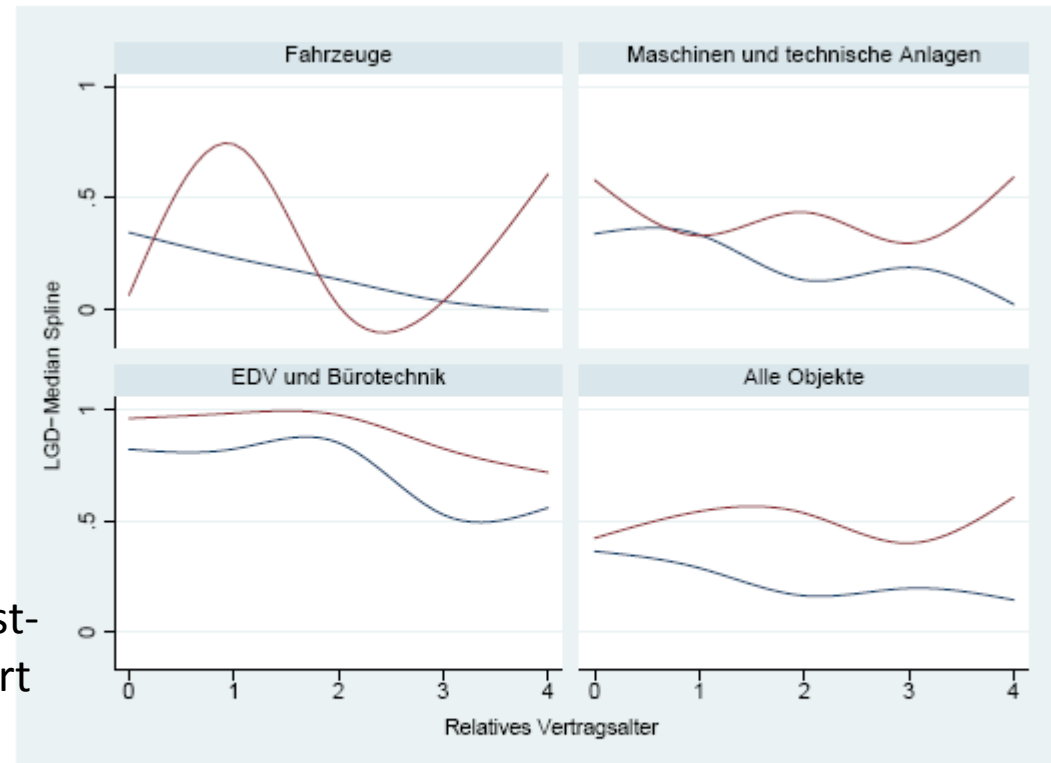
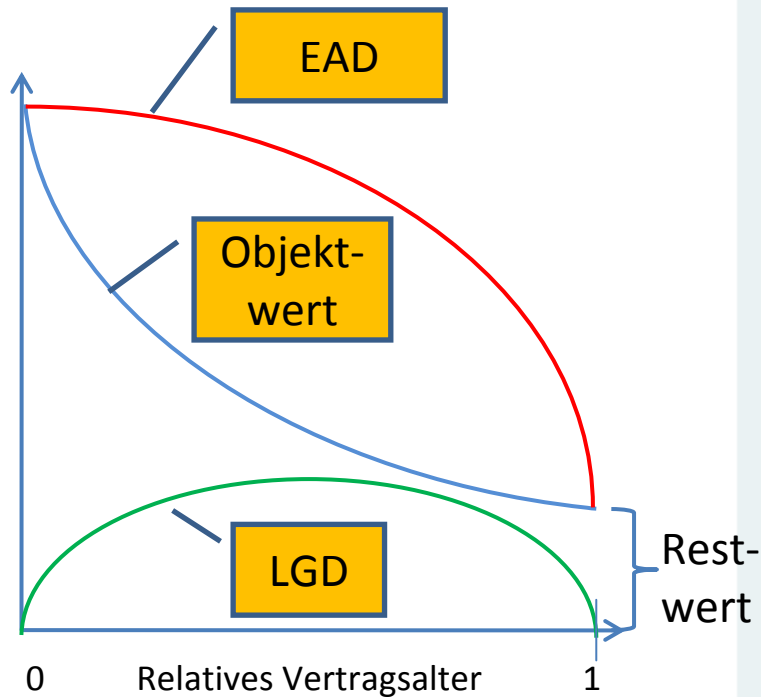


Abbildung 12.10: Median-Bänder der Verlustquote abgewickelter Verträge in Abhängigkeit des relativen Vertragsalters und getrennt nach einem Workout mit Objektverwertung (blau) und ohne Objektverwertung (rot).

LGD-Schätzung


- Über alle Objektkategorien, Kundenarten, Vertragsarten hinweg unterscheiden sich die LGDs wiedergesundeter Verträge deutlich von denen abgewickelter Verträge
- Getrennte Schätzung der LGDs für wiedergesundete und abgewickelte Verträge

$$\text{LGD}^* = p_G \cdot \text{LGD}_G^* + p_A \cdot \text{LGD}_A^*$$

1. Verwendung historischer Durchschnittswerte (Mittelwert/Median) als Schätzwerte für p_G , LGD_G^* und LGD_A^* ? **oder**
2. Identifizierung von Faktoren, die den LGD_G , LGD_A und p_G beeinflussen; Schätzung von LGD_G , LGD_A und p_G in Abhängigkeit von der Ausprägung dieser Faktoren

Schätzprobleme (1)

Problem: nicht alle Faktoren, die auf die Höhe von LGD_G , LGD_A und p_G Einfluss haben, sind zu Vertragsbeginn bekannt!



Vertragsbeginn

- Objektmerkmale
 - Objektart
 - neu/gebraucht
- Kundenmerkmale
 - Gewerblich
 - Privat
 - Branche
- Vertragsmerkmale
 - Vereinb. Laufzeit
 - Kalk. Restwert
 - Optionsrechte
 - Garant. Restwert
 - Mietvorauszahlung

Ausfallzeitpunkt

- Ausfallgrund
- Ausfalldatum
- relatives Vertragsalter bei Ausfall
- Wert des Objekts
- Ausfallrate



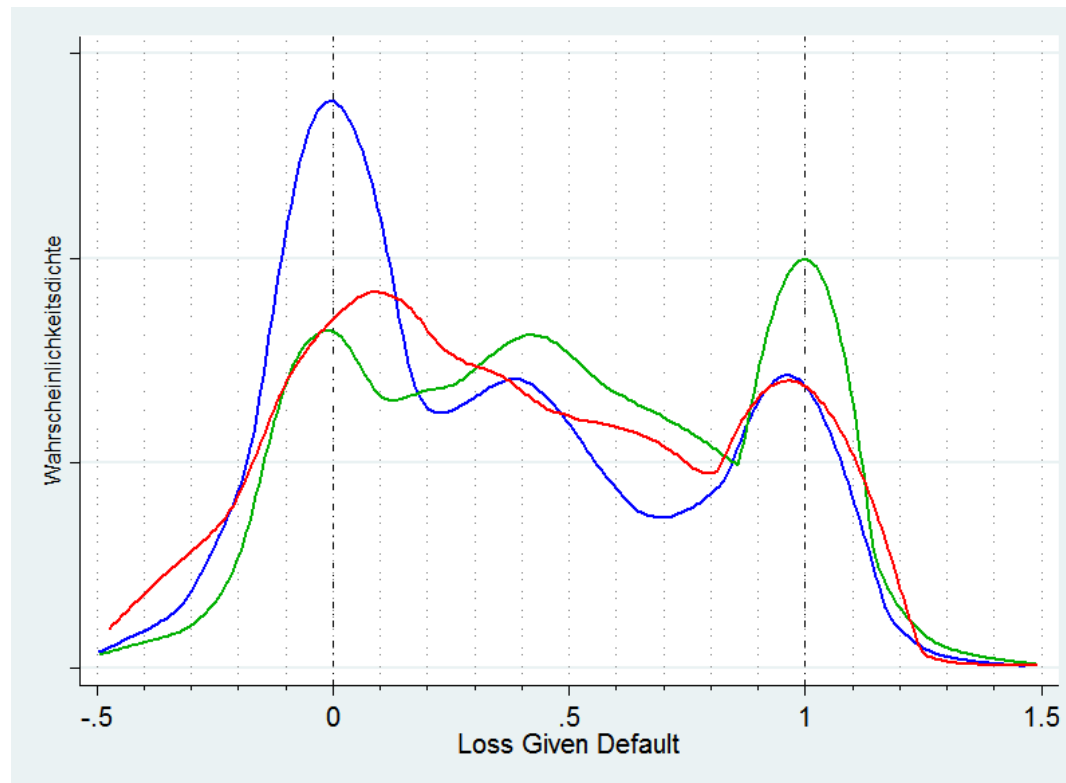
Update erforderlich
§ 132, 9 SolvV

Ende des Workouts

- Abwicklung/Wiedergesundung
- Verwertung/keine Verwertung
- Dauer des Workouts

Schätzprobleme (2)

Wie geht man mit der multi-modalen Form der Wahrscheinlichkeitsdichte der Verlustquote um?



Schätzprobleme (3)

- Lösung hängt ab von
 - der Datenverfügbarkeit/Datenqualität
 - Schätzaufwand, der betrieben werden soll
 - der Möglichkeit, Faktoren zu identifizieren, die den LGD systematisch beeinflussen
 - Leistungsfähigkeit der mathematisch/statistischen Verfahren
- Lösungsansatz
 - p_G : historische Durchschnittswerte (evtl. getrennt nach Objektkategorien)
 - LGD_A : LGD-Schätzung auf der Basis identifizierter Einflussgrößen; Mischverteilungsansatz zur Berücksichtigung der multimodalen Form der LGD-Verteilung
 - LGD_G : historische Durchschnittswerte

LGD_A-Schätzung mit einem Mischverteilungsansatz

Bestimmung von Vertragsklassen homogener Verlustquoten unter Verwendung der zum Ausfallzeitpunkt und am Ende des Workouts verfügbaren Informationen

finite mixture of regression model; funktionale Beziehung zwischen dem LGD und erklärenden Variablen



Zuordnung der Verträge zu den Klassen unter Verwendung der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Logistische Regression, Nächste-Nachbarn-Methode



Schätzung klassenspezifischer Verlustquoten auf der Basis der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Regressionsanalyse

LGD_A-Schätzung mit einem Mischverteilungsansatz

Bestimmung von Vertragsklassen homogener Verlustquoten unter Verwendung der zum Ausfallzeitpunkt und am Ende des Workouts verfügbaren Informationen

finite mixture of regression model; funktionale Beziehung zwischen dem LGD und erklärenden Variablen



Zuordnung der Verträge zu den Klassen unter Verwendung der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Logistische Regression, Nächste-Nachbarn-Methode



Schätzung klassenspezifischer Verlustquoten auf der Basis der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Regressionsanalyse

Schätzmethodik

Bestimme

- aus wie vielen Klassen (Verteilungen) sich die Gesamtverteilung zusammensetzt (N)
- den Typ der Verteilungen (Normalverteilung, Exponentialverteilung, ..., identischer Verteilungstyp?)
- die Parameter der Verteilungen (z.B. Erwartungswert, Varianz)
- die Gewichte π
- von welchen Faktoren ($\beta \cdot X$) die Zuordnung zu den Klassen abhängt, und wie der Einfluss ist (Werte für β)
- die Wahrscheinlichkeit, mit der ein Vertrag j einer Klasse k zugeordnet wird (p_{jk})

Lösungsmethode: Expectation/Maximization-Algorithmus

Zuordnung hängt auch von Leasinggeber-spezifischen Faktoren ab!

Unabhängige Variablen	$\mathcal{K} = 1$	$\mathcal{K} = 2$		$\mathcal{K} = 3$		
	$k = 1$	$k = 1$	$k = 2$	$k = 1$	$k = 2$	$k = 3$
LG B	0.146***	0.003	0.069***	0.018*	0.127***	0.027**
LG C	0.120***	0.086***	0.090***	0.089***	0.140***	0.010
Workout	0.001***	-0.001	0.001**	-0.001	0.001*	0.001***
Int(LG B,W)	0.001***	0.001***	0.001***	0.001***	0.001***	-0.001**
Int(LG C,W)	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001
Rel. Alter	-0.302***	-0.080***	-0.235***	-0.071**	-0.182***	-0.087***
Rel. Alter 2	0.102***	0.018	0.073***	0.014	0.058***	0.028**
Objekterlös	-0.443***	-0.041***	-0.444***	-0.077***	-0.557***	-0.167***
ln(EAD)	-0.048***	0.001	-0.063***	0.001	-0.047***	-0.031***
Konstante	1.259***	0.048	1.625***	0.061*	1.396***	1.329***
π_k	1	0.32	0.68	0.31	0.49	0.20
AIC	8807	6566		5509		
BIC	8880	6734		5765		

Qualität der probabilistischen Zuordnung

Statistische Kennziffern	$\mathcal{K} = 1$	$\mathcal{K} = 2$		$\mathcal{K} = 3$		
	$k = 1$	$k = 1$	$k = 2$	$k = 1$	$k = 2$	$k = 3$
Mittelwert (μ_k)	0.44	-0.04	0.65	-0.06	0.76	0.43
Standardabweichung (σ_k)	0.42	0.14	0.31	0.14	0.23	0.36
5%-Perzentil (Q_5)	-0.15	-0.30	0.17	-0.33	0.40	0.01
25%-Perzentil (Q_{25})	0.05	-0.10	0.39	-0.13	0.56	0.16
50%-Perzentil (Q_{50})	0.41	-0.02	0.63	-0.05	0.74	0.31
75%-Perzentil (Q_{75})	0.82	0.04	0.98	0.00	0.99	0.70
95%-Perzentil (Q_{95})	1.05	0.16	1.08	0.14	1.11	1.04

Modell	Klasse	Mittelwert	Median	Std.abw.	25%-Perzentil
$\mathcal{K} = 2$	$k = 1$	0.85	0.90	0.15	0.73
	$k = 2$	0.95	1.00	0.11	0.98
$\mathcal{K} = 3$	$k = 1$	0.84	0.90	0.16	0.71
	$k = 2$	0.81	0.87	0.19	0.62
	$k = 3$	0.65	0.63	0.12	0.56

Institutsspezifische Klassenzuordnung (LG A)

Unabhängige Variablen	$\mathcal{K} = 1$	$\mathcal{K} = 2$	
	$k = 1$	$k = 1$	$k = 2$
Workout	0.001***	-0.001**	0.001***
Rel. Alter	-0.404***	-0.152***	-0.391***
Rel. Alter 2	0.146***	0.054***	0.142***
Objekterlös	-0.192***	0.022*	-0.227***
ln(EAD)	-0.078***	-0.006	-0.097***
Retail	-0.030**	-0.001	-0.022
Anzahlung	-0.036**	0.001	-0.015
Mietkauf	-0.291***	0.017	-0.370***
Sicherheit	-0.122***	0.014	-0.165***
Konstante	1.514***	0.092*	1.905***
π_k	1	0.35	0.65
AIC	2253	1417	
BIC	2319	1566	

Qualität der probabilistischen Zuordnung

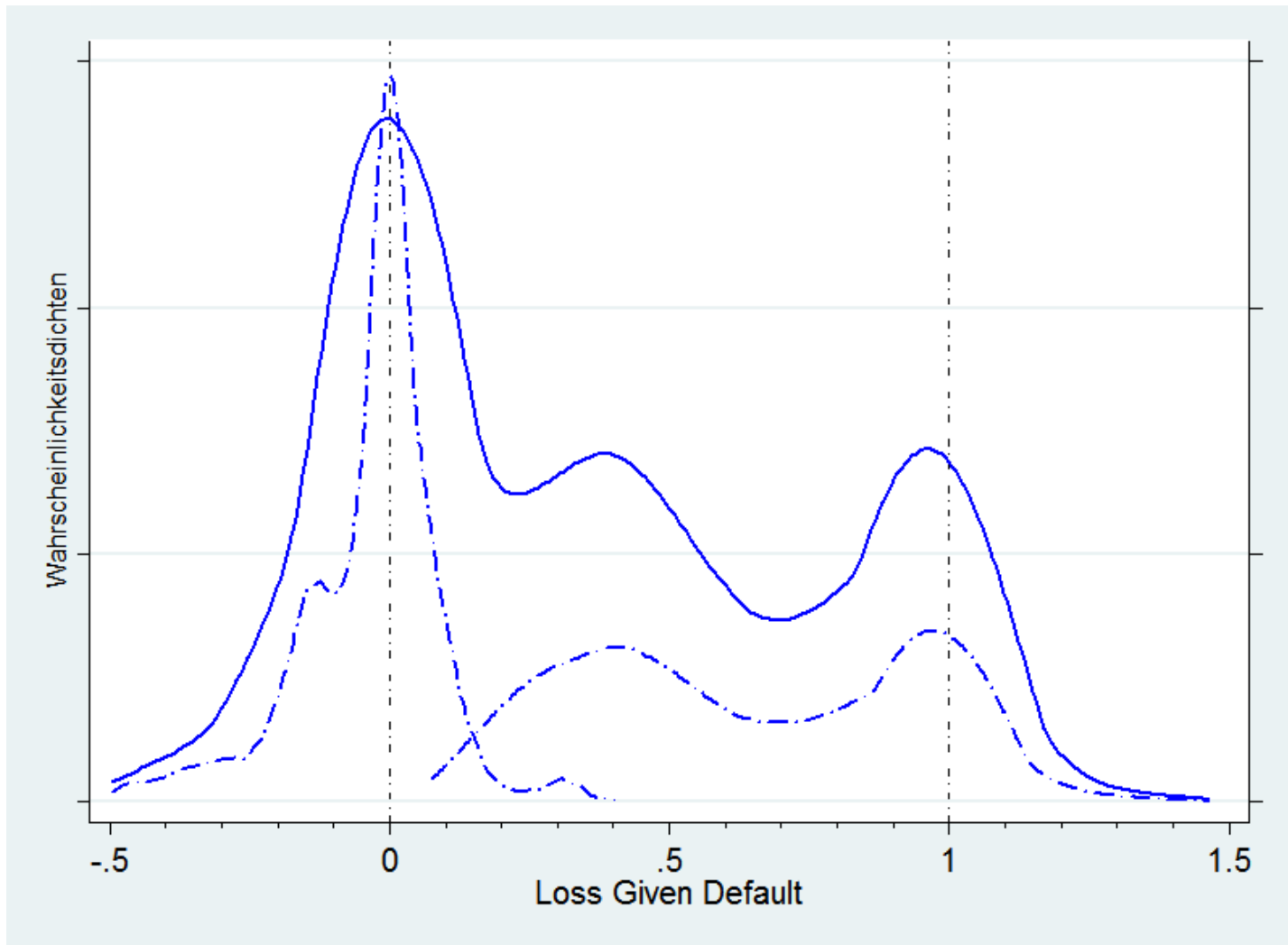
Institut A – Abschätzung der Qualität der verteilungsbasierten Clusteranalyse anhand der Lage- und Streuungsparameter der bedingten Verteilungen und der Verteilungen der a-posteriori Wahrscheinlichkeiten für die jeweilige Klassenzugehörigkeit

Statistische Kennziffern	$\mathcal{K} = 1$	$\mathcal{K} = 2$	
	$k = 1$	$k = 1$	$k = 2$
Mittelwert (μ_k)	0.35	-0.02	0.59
Standardabweichung (σ_k)	0.41	0.15	0.35
5%-Perzentil (Q_5)	-0.18	-0.23	0.08
25%-Perzentil (Q_{25})	0.00	-0.07	0.34
50%-Perzentil (Q_{50})	0.27	0.00	0.55
75%-Perzentil (Q_{75})	0.70	0.03	0.93
95%-Perzentil (Q_{95})	1.02	0.14	1.06

Klasse	Mittelwert	Median	Std.abw.	25%-Perzentil
$k = 1$	0.77	0.77	0.15	0.64
$k = 2$	0.95	1.00	0.12	0.98

Mediane der a-posteriori Wahrscheinlichkeiten liegen für die Institute B und C bei 97% (79%) und 99% (100%) für die Klasse 1 (2)

LGD-Verteilung der beiden Cluster



LGD_A-Schätzung mit einem Mischverteilungsansatz

Bestimmung von Vertragsklassen homogener Verlustquoten unter Verwendung der zum Ausfallzeitpunkt und am Ende des Workouts verfügbaren Informationen

finite mixture of regression model; funktionale Beziehung zwischen dem LGD und erklärenden Variablen



Zuordnung der Verträge zu den Klassen unter Verwendung der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Logistische Regression, Nächste-Nachbarn-Methode



Schätzung klassenspezifischer Verlustquoten auf der Basis der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Regressionsanalyse

Qualität der Zuordnung

- Ergebnisse einer Klassifikation mit drei bzw. fünf nächsten Nachbarn auf Basis der zu Vertragsbeginn verfügbaren Informationen:

Gesellschaft	Schätz- und Ist-Klassen	Klassifikation ($k = 3$)		Klassifikation ($k = 5$)		Total
		1	2	1	2	
A	Ist-Klasse 1	70.78%	29.22%	60.77%	39.23%	100%
	Ist-Klasse 2	13.47%	86.53%	14.39%	85.61%	100%
	Total ($\hat{\pi}_k$)	36.81%	63.19%	33.28%	66.72%	100%
	Gesamtfehler	0.213		0.268		
B	Ist-Klasse 1	65.72%	34.28%	79.54%	20.46%	100%
	Ist-Klasse 2	0.97%	99.03%	32.36%	67.64%	100%
	Total ($\hat{\pi}_k$)	51.75%	48.25%	69.36%	30.64%	100%
	Gesamtfehler	0.176		0.264		
C	Ist-Klasse 1	80.84%	19.16%	75.05%	24.95%	100%
	Ist-Klasse 2	26.34%	73.66%	31.92%	68.08%	100%
	Total ($\hat{\pi}_k$)	55.11%	44.89%	54.69%	45.31%	100%
	Gesamtfehler	0.228		0.284		

- Klassifikation bei Ausfall führt zu geringeren Gesamtfehlern (19.9%, 15.9% bzw. 19.8% im Falle von $k = 3$ nächsten Nachbarn)

LGD_A-Schätzung mit einem Mischverteilungsansatz

Bestimmung von Vertragsklassen homogener Verlustquoten unter Verwendung der zum Ausfallzeitpunkt und am Ende des Workouts verfügbaren Informationen

finite mixture of regression model; funktionale Beziehung zwischen dem LGD und erklärenden Variablen



Zuordnung der Verträge zu den Klassen unter Verwendung der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Logistische Regression, Nächste-Nachbarn-Methode



Schätzung klassenspezifischer Verlustquoten auf der Basis der zu Vertragsbeginn bzw. Ausfall verfügbaren Informationen

Regressionsanalyse

Schätzung der Verlustquote innerhalb der einzelnen Klassen

Variablen	Vertragsabschluss			Schuldnerausfall		
	$\mathcal{K} = 1$	$\mathcal{K} = 2$		$\mathcal{K} = 1$	$\mathcal{K} = 2$	
	$k = 1$	$k = 1$	$k = 2$	$k = 1$	$k = 1$	$k = 2$
Retail	-0.037***	0.017	-0.040**	-0.042***	0.006	-0.042**
Anzahlung	0.001	0.014	-0.008	0.010	0.021	-0.006
Mietkauf	-0.056**	-0.009	-0.182***	-0.038	0.004	-0.175***
Sicherheit	-0.050***	0.004	-0.045**	-0.058***	-0.007	-0.049**
Fahrzeuge	-0.153***	-0.080**	-0.043	-0.147***	-0.085**	-0.021
Maschinen	-0.144***	-0.062**	-0.030	-0.149***	-0.062**	-0.040
Infrastruktur	0.255	-0.121	0.356*	0.255	-0.141	0.380*
EDV	0.275***	0.143***	0.268***	0.291***	0.130***	0.288***
ln(AW)	-0.040***	-0.026***	-0.035***	-0.103***	-0.071***	-0.091***
Gebraucht	-0.003	0.048	-0.010	-0.012	0.028	-0.019
Anschluss	0.084***	0.047	0.049*	0.126***	0.056	0.101***
Laufzeit	0.015***	0.018***	0.014*	0.020**	0.023**	0.025**
Rel. Alter	-	-	-	0.028	-0.043	0.107
Rel. Alter 2	-	-	-	0.011	0.073	-0.017
ln(EAD)	-	-	-	0.074***	0.046**	0.071***
Konstante	1.369***	1.344***	0.701*	1.111***	0.617	0.435
Adj. R ²	33%	17%	26%	35%	22%	27%

Vertragsmerk-
male

Objektmerk-
male

Kundenmerk-
male

Validierung

Mittlerer quadratischer Fehler

$$\text{mse} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{LGD} - \text{LGD}^*)^2}$$

Mittlerer absoluter Fehler

$$\text{mad} = \frac{1}{n} \sum_{i=1}^n |\text{LGD} - \text{LGD}^*|$$

1. Median pro Merkmalskombination aus Objektkategorie/Wirtschaftsbereich
 2. Lineare Regression ohne Klassenunterschiede
 3. 2-Klassen-Mischverteilungsansatz; Zuordnung auf die Klassen mit der 3-nächste-Nachbarn-Methode; lineare Regression pro Klasse
- Mit Zensierung bei $\text{LGD} = 0$
 - Ohne Zensierung

Validierung der Schätzungen

Zensur	Gesellschaft	Modell	mad		mse	
			Beginn	Ausfall	Beginn	Ausfall
Nein	A	1	0.258	0.258	0.359	0.359
		2	0.263	0.259	0.336	0.331
		3	0.229	0.226	0.310	0.306
	B	1	0.330	0.330	0.405	0.405
		2	0.319	0.314	0.386	0.382
		3	0.303	0.294	0.376	0.368
	C	1	0.326	0.326	0.408	0.408
		2	0.315	0.310	0.379	0.374
		3	0.264	0.263	0.340	0.339
Ja	A	1	0.232	0.232	0.330	0.330
		2	0.237	0.232	0.306	0.302
		3	0.203	0.199	0.281	0.279
	B	1	0.311	0.311	0.377	0.377
		2	0.300	0.295	0.358	0.355
		3	0.284	0.275	0.348	0.342
	C	1	0.295	0.295	0.372	0.372
		2	0.283	0.278	0.344	0.340
		3	0.233	0.233	0.309	0.309

Offene Punkte

- Stabilität der Ergebnisse über die Zeit
- Weitere Einflußgrößen berücksichtigen
- Downturn-LGD's