

大规模同质不良资产组合回收率定价研究^①

马宇超, 陈暮紫, 陈浩, 王博, 陈敏, 杨晓光

(中国科学院数学与系统科学研究院, 北京 100190)

摘要: 处置不良资产是我国金融业改革和发展的重要问题, 大规模批量处置不良贷款是处置不良资产的首选方法之一. 不良贷款组合的回收率分布, 是不良贷款组合定价的基础. 针对不良贷款回收率的双峰分布特性, 将不良贷款分为低峰回收贷款和高峰回收贷款, 证明了不良贷款组合的回收率收敛于低峰回收率和高峰回收率的条件期望之和. 基于信用风险结构模型, 进一步证明了高峰回收率条件期望的正态逆变换与低峰回收比率的正态逆变换之间存在线性关系. 由于对低峰回收贷款易于判断, 很容易估计低峰回收比率, 因而可以通过线性回归估计出高峰回收率条件期望, 这样就给出了不良贷款组合的整体回收率估计模型. 基于这一模型, 还给出了估计不良贷款回收率的分位数的计算方法, 该方法实际上是 VaR 方法的推广.

关键词: 不良贷款组合; 巨额同质资产组合; 回收率估计; 结构模型

中图分类号: F830 **文献标识码:** A **文章编号:** 1007-9807(2010)02-0086-11

0 引言

处置不良资产是我国金融业改革和发展的核心问题. 我国的 4 大资产管理公司肩负着化解不良资产风险、维护金融体系稳定的重任. 在 2006 年底之前, 金融不良资产主要是指随着 4 大国有银行准备上市而政策剥离给 4 大资产管理公司的不良贷款. 随着各个城市商业银行的上市, 商业化收购不良资产的手段完善, 近年来, 通过商业化竞购的不良资产也成为国内金融不良资产的重要组成部分. 截至 2007 年 3 季度末, 中国商业银行仍持有 12 517.8 亿不良贷款. 中国 4 大资产管理公司于 2006 年底处置完结 1.4 万亿元政策性剥离的不良资产后, 未处置不良资产存量仍然巨大, 其中绝大部分是不良贷款. 如何对规模庞大, 质量庞杂的不良贷款组合进行有效的风险定价、管理和预警, 是亟需解决的问题.

国际上, 针对违约贷款的信用风险和违约损

失率的监管和风险管理已经纳入巴塞尔新资本协议 (Basel II) 中. 巴塞尔新资本协议 (Basel II) 在内部评级法 (RB) 中, 提出用违约概率 (PD)、违约损失率 (LGD)、违约风险暴露 (EAD) 和有效期限 (M) 等风险要素来确定风险资本要求. 在高级法中允许银行自行估计相关要素. 随着风险管理技术水平的提高, 高级内部评级法成为国际上大银行的首选目标. 因此, 实现对内部评级法中风险要素的准确估值已经成为各国银行业所关注的重要问题. 其中违约贷款的损失率 (LGD) 的估值是高级内部评级法实施中的重点和难点.

国外针对 LGD 的研究主要基于对违约债券的统计分析, 这是鉴于违约债券的公开数据较多的原因. 如 Bahar 和 Brand (1998) 比较了不同优先级债券的回收率差异, Altman 和 Kishore (1996) 则广泛分析了不同行业的违约债务回收率特征, Aller 和 Saunders (2002) 探讨了是否应该在信用风险模型中引入宏观经济周期变量. Hu 和

① 收稿日期: 2008-05-12; 修订日期: 2009-09-11.

基金项目: 国家基础研究计划资助项目 (973 项目) (2007CB814902); 国家自然科学基金创新研究群体资助项目 (70221001); 国家自然科学基金重点资助项目 (70331001); 国家自然科学基金海外杰出青年基金资助项目 (10628104).

作者简介: 马宇超 (1983—), 男, 安徽合肥人, 博士. Email: mayucao@ms.ac.cn

Perraudin^[1]在前述研究中所考虑的因素基础上, 用穆迪的数据库建立了预测回收率的简单线性回归模型. 由于国内企业债市场很小, 国内对于 LGD 的研究主要是针对不良贷款. 由于贷款数据的限制和不公开性, 针对不良贷款的 LGD 模型研究和实证研究相当有限. 王春峰等^[2]将判别分析法应用于商业银行信用风险评估中, 并且通过与 logit 方法相比较, 进一步研究了判别分析法的有效性. 最后讨论判别分析法作为传统的建模工具的优缺点. 梁世栋等^[3]构造了信用风险期限结构的框架性模型, 属于强度模型流派, 并讨论了两因子模型的例子, 给出了可违约债券的价格解析表示式, 最后分析了信用风险衍生产品的定价问题. 石晓军等^[4]在对边界 Logistic 违约率模型展开进一步研究, 利用中国公司数据展开实证研究, 找到了 Cramer 问题的中国证据. 然而无论是国际国内, 对于 LGD 预测研究的出发点大多都是对针对单笔债项, 对于大规模不良贷款组合资产 LGD 的研究尚未有发现.

针对大规模信用资产 (贷款或债券组合) 的定价模型可以依靠 Copula 方法 (David Li (2000)) 和 LHP (large homogenous portfolio) 逼近方法 (Vasicek (1987)) 来度量违约相依性. 在处理资产池中各种资产质量参差不齐的条件下, Copula 方法能够很好的度量其相依性. 但是 Copula 方法的劣势在于依靠 Copula 函数估计联合边际分布, 在资产池相当庞大时 Copula 方法往往舍本逐末, 过于关注单笔贷款之间的相关性, 容易过度拟合 (over fit). 此外, Copula 方法也容易低估系统性风险和流动性风险. 而面对中国 4 大资产管理公司数十亿乃至上千亿的不良资产包, 违约贷款的笔数以万计, 在巨额不良贷款组合的定价和管理中, IHP 逼近方法的思想更为直观和合理.

Vasicek (1987) 年就提出 IHP 模型的思想, 但是直到 Vasicek 的文献 [5] 的发表, 才标志着 IHP 模型的成熟和理论上的完善. 由于正态分布在描述重尾现象的缺陷, Schlegel (2005) 给出了 IHP 方法推广到 t 分布和 t -Copula 的情形. Martin^[6]也是基于对违约概率和违约损失率的联合度量的基础上, 给出了如何对基于独立假定的正常贷款组合违约损失率进行建模和估计.

但是纵观所有文献, 目前的违约概率和违约

损失率相关模型和 IHP 的信贷资产组合定价方法并不能直接应用于不良贷款组合估算中. Vasicek^[5]在考虑贷款组合的价值过程中, 是从银行的观点考虑贷款资产组合价值, 包括违约概率 (PD) 和违约损失率 (LGD) 的联合分布, 而对于贷款资产组合价值起主导作用的是 PD, PD 基本上决定着整个资产组合的价值, 因而 Vasicek 的模型中关于 LGD 假设就相当简单. 而从资产管理公司的角度来看, 中国 4 大资产管理公司管理的资产全部都为违约贷款, 不存在估计 PD 的问题, 更不需要考虑 PD 和 LGD 的关系, LGD 决定着不良贷款的全部价值, 需要的是给出更为精细的 LGD 估计方法, 从而为不良资产竞购、处置、打包证券化和风险控制提供参考.

本文试图在不良贷款组合的 LGD 研究方面做一些探索, 研究的意义在于: 第 1, 以往的研究更多是注重单个资产的回收率建模, 从资产组合角度进行回收率建模的研究比较少. 第 2 对于不良贷款的处理, 资产管理公司关注的非常重要的处置方式是打包处置, 不良贷款资产组合的回收率研究有着很大的现实意义. 如何快速估算资产包的价值, 在商业化不良资产收购中是非常重要的一个问题. 由于不良资产商业化收购, 标的金额巨大, 往往涉及数百亿乃至上千亿, 倘若能够准确快速的估算不良资产组合回收率, 对资产管理公司意义巨大. 第 3 资产管理公司需要从全辖以及分支机构的角度进行业绩考评、绩效评估和风险管理, 需要以资产组合为对象计算回收率.

本文针对不良贷款组合的回收率估计模型的建立, 基于 3 个基础的前提假设: 1) 根据不良贷款回收率的双峰性质, 把不良贷款分为低峰回收贷款和高峰回收贷款两类. 2) 低峰回收贷款易于判别, 在真正的回收行为开始之前, 即可以根据其贷款信息、公司基本属性和基本财务信息得以判别. 在实际的回收过程中, 抵押担保是否有效, 公司规模等因素直接影响最终回收率, 这也是 Logistic 回归、支持向量机、Bayes 判别等判别分析模型得以广泛使用, 且效果尚佳的原因. 3) 影响整个不良贷款组合回收率的系统风险因子是时变的, 这点可以从图 3 得到实证. 假设系统风险因子只影响高峰回收率, 低峰回收率是相对不变的.

模型的基本思路是整个资产组合的回收率收敛于低峰回收率和高峰回收率的条件期望之和,而高峰回收率又与低峰回收比率之间存在一定的关系,可以通过低峰回收比率计算出高峰回收率,从而得到对不同时间段整个不良贷款组合的违约回收率的估计。

模型的优点首先是计算实现性比较好。由于用于单笔贷款的判别类模型易于估计低峰回收贷款,由此可以将用于单笔贷款的判别类模型和不良贷款组合的模型联系起来,解决了目前模型难于分析不良贷款组合回收率的问题。其次,模型中将系统风险因子纳入考虑的范围,可以得到不同宏观经济条件下不良贷款组合的回收率。这也是国内首次在研究不良贷款组合时将时变系统风险因子纳入建模框架。

本文的架构如下,第 2 节叙述大规模同质不良贷款组合回收率的估计方法。其中,1.1 节描述使用 Bernoulli 变量构造不良贷款违约回收率的方法;1.2 节证明随着违约贷款组合规模的不断扩大,在各笔违约贷款异质风险独立的假设条件下,违约贷款组合的总体回收率收敛于条件期望,由此说明使用高峰回收率期望作为不良贷款组合整体回收率估计的合理性;1.3 节阐述低峰回收率的估计方法和统计性质;1.4 节阐述利用低峰回收比率估计高峰回收率条件期望的建模方法,并给出实证结果说明这种方法的合理性。第 3 节给出不良贷款组合的回收率分位数估计,并解释其分位数在不良贷款风险管理中的作用。第 4 节叙述本文的结论和政策建议。

1 大规模同质不良贷款组合回收率估计方法

1.1 基于 Bernoulli 变量不良贷款回收率模型

LOSSMETRICS^②是我国最大的不良贷款清收数据库。对该数据库的不良贷款违约回收率进行统计检验,发现在违约贷款的回收率的经验分布中,呈现明显的双峰分布,即回收率大多聚集于

0 和 1 附近(如图 1 所示)。国外对 LGD 的诸多研究也表明,违约回收率的分布可以用 Beta 分布来刻画,呈现标准的双峰特征。

因此对于大规模同质的不良贷款,可以假设它们的违约回收率服从 Bernoulli 变量:即分为“低峰回收率”和“高峰回收率”两种情形。更详细的数据分析(陈暮紫等)表明,高峰回收率呈现时变的形态(如图 2 数据来源 LOSSMETRICS),而低峰回收率基本上保持为不变的常数。本文认为高峰回收率这种时变特性,是受到宏观经济等因素带来的系统风险因子造成的,而低峰回收率由于接近于 0 和系统风险因子并不呈现出明显的相关性,因而保持常态。

用 Y 表示代表系统风险因子的随机变量,记 α 为低峰回收率, $\beta(Y)$ 为高峰回收率, $R(Y)$ 为依赖于系统风险因子 Y 的回收率,则

$$R(Y) = \begin{cases} \alpha & P(R = \alpha) = PL(Y) \\ \beta(Y) & P(R = \beta(Y)) = 1 - PL(Y) \end{cases} \quad (1)$$

其中 $PL(Y)$ 可以定义为“低峰回收概率”。

定义“低峰回收概率”(probability of low-recovery) 为

$$PL(Y) = P\{LGD = 1 - \alpha \mid Y\} \\ = P\{Recovery = \alpha \mid Y\}$$

“高峰回收概率”(probability of high-recovery) 为

$$PH(Y) = 1 - P\{LGD = 1 - \beta \mid Y\} \\ = P\{Recovery = \beta \mid Y\}$$

对国内的数据实证研究(陈浩等(2007))表明, α 在 0.05 附近, $\beta(Y)$ 均值在 0.4 附近。在模型的建立中,假设资产管理公司拥有大规模的同质贷款^③,即所有贷款为同质资产。所谓同质资产,是指不良资产组合满足以下条件:1)所有贷款都承担同样的系统风险,2)所有贷款的违约回收率都服从同一概率分布,3)假设这些不良贷款之间的异质风险是互相独立的。

②注: LossMetrics 数据库是东方资产管理公司旗下的东和数据公司建立的目前国内最大的中国银行业违约债务损失数据库。

③Vasicek(1987)年就提出 IHP 建模的思想, Vasicek 文献[5]的发表,标志着 IHP 模型的成熟和理论上的完善。对其做技术上的修改,给出不同的 Bernoulli 变量,可以用以衡量不良贷款组合的风险。

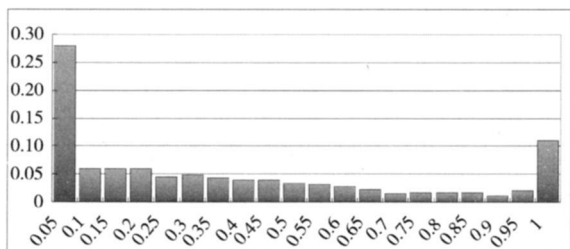


图 1 违约回收率直方图

Fig. 1 Histogram of loss given default

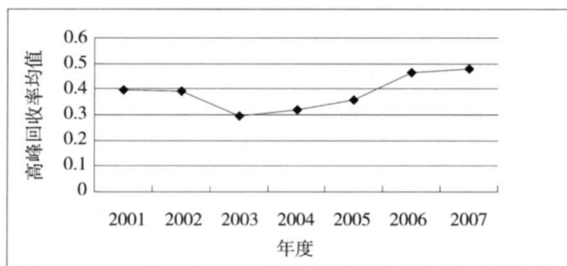


图 2 高峰回收率均值的时变特性

Fig. 2 Time-dependent feature of high-recovery LGD

下面引入本文最重要的一个概念: 低峰回收比率. 对于资产管理公司的违约贷款, 其“低峰回收比率”是指给定一个不良贷款资产包, 其中回收率为低峰回收率的贷款个数与贷款组合总数之比, 其数学表述如下:

某个违约贷款组合, 共有 n 笔违约贷款, 定义第 i 笔违约贷款违约回收率为低峰 (即回收率为 α) 的示性变量为

$$D_i = \mathbb{I}_{\{Recovery_i = \alpha\}}$$

则“低峰回收比率” D 为

$$D = \frac{\sum_{i=1}^m D_i}{m} \quad (2)$$

容易看出, 取定样本区间的“低峰回收比率” D 是 $PL(Y)$ 的自然估计. 对于低峰回收比率, 一个粗糙的办法是可以“用样本低峰回收比率”来估计, 即通过给定低峰限定值 (如 0% 或 5%), 计算样本中回收率低于 α 的样本比率, 作为“低峰回收比率”的估计. 此外, 在贷款信息和借款公司财务指标足够的情况下, 低峰回收比率易于预测, 有众多的判别分析方法可以用来预测低峰回收比率. 通过 LossMetrics 数据库训练样本, 使用 Logistic、传统马氏距离判别方法、Bayes 判别、支持向量机、神经网络等各种方法 (参见唐跃 (2007)), 均能较好的在不良贷款回收之前, 判别出该笔贷款是否回收率较低 (或基本无回收), 从而易于得到

不良贷款组合中的“低峰回收比率”.

1.2 大规模同质不良贷款组合回收率的收敛性

目前, 国内各个资产管理公司的不良贷款规模以千亿计量, 其规模仍在不断扩大. 如此规模庞大的不良贷款组合的回收率能否被预测以及如何预测是非常重要的课题. 在模型中, 本文定义了基于系统风险因子 Y 的“高峰回收率”的条件期望

$$\beta(Y) = \mathbb{E}[Recovery | D = 0, Y]$$

($D = 0$ 表示 $\beta(Y)$ 为基于所有非低峰回收情形的均值), 以此来预测不良贷款组合. 很自然的问题是: 这种预测不良贷款组合回收率的方法是否合理?

事实上本文证明: 随着违约贷款组合规模的不断扩大, 在各笔违约贷款异质风险独立的假设条件下, 违约贷款组合的总体回收率收敛于条件期望. 当衡量对巨额违约贷款组合的风险和定价时, 这个结论尤为重要. 以下将证明这个定理.

在证明定理之前, 需要定义条件期望和相关变量. 回收统计量 Bernoulli 变量 R_i 可以改写成

$$R_i = \mathbb{I}_{\{P|V_i < K\}} \times \alpha + \mathbb{I}_{\{P|V_i \geq K\}} \times \beta(Y) \quad (3)$$

其期望为

$$\mathbb{E}(R_i) = \alpha \times PL(Y)_i + \beta(Y) \times (1 - PL(Y)_i)$$

定义整个不良贷款组合的总体回收率 $R^{(m)}$

$$\begin{aligned} R^{(m)} &= \sum_{i=1}^m \omega_i R_i \\ &= \alpha \sum_{i=1}^m \omega_i \mathbb{I}_{\{r_i < s\}} + \beta \sum_{i=1}^m \omega_i \mathbb{I}_{\{r_i > s\}} \end{aligned} \quad (4)$$

令 ω_i 为第 i 笔不良贷款的权重

$$\omega_i = \frac{N_i}{\sum_{i=1}^m N_i}$$

N_i 为第 i 笔不良贷款的债权本息合计金额. 对任意 $y \in \mathbb{R}$, 第 i 笔违约贷款的条件期望满足

$$\begin{aligned} \mathbb{E}(R_i | Y = y) &= \mathbb{P}(D_i = 1 | Y = y) \times \\ &\mathbb{E}(R_i | D_i = 1, Y = y) + \mathbb{P}(D_i = 0 | Y = y) \times \\ &\mathbb{E}(R_i | D_i = 0, Y = y) \\ &= PL_i(y) \times \alpha + PH_i(y) \times \beta(y) \\ &= PL_i(y) \times \alpha + (1 - PH_i(y)) \times \beta(y) \end{aligned} \quad (5)$$

由于上式对任意 $y \in \mathbb{R}$ 和 $0 < i \leq m$ 成立, 故可以

$$\mathbb{E}(R^{(m)} | Y) = PL(Y) \times \alpha + (1 - PL(Y)) \times \beta(Y) \quad (6)$$

其中, $PL(Y) = N(c - eY)$ 将在式(14)中给出, 为单因子结构模型下的表达式.

给出以下定理以保证条件期望的合理性.

定理 1 考虑巨额同质资产组合, $R^{(m)}$ 如上定义, 则 $R^{(m)}$ 满足下式

$$P(\lim_{m \rightarrow \infty} [R^{(m)} - \mathbb{E}(R^{(m)} | Y)] = 0) = 1 \quad (7)$$

证明见附录.

尽管以上定理保证了条件期望可以用来估计不良贷款组合价值并为其定价, 但是从式(6)可以看出如何计算 $\mathbb{E}(R^{(m)} | Y)$ 仍然有困难, 因为 $\beta(Y)$ 和 $PL(Y)$ 并不互相独立.

Martin^[6] 提出一个新的模型用于解决违约概率和违约损失率联合分布的问题, 将 Y 的系统风险重新组合, 构造新的系统风险因子 Z , 从而将变量分离违约回收率和违约概率分离. 借鉴其建模技术, 可以解决 $\beta(Y)$ 和 $PL(Y)$ 并不互相独立的问题. 此种方法减少了对数据量的需求, 从而仅仅有较少的参数需要估计, 而在系统风险因子 Z 中可以包含行业、地区等各种系统风险.

实际上, 在不良贷款组合的模型中, Y 可以看做多个系统风险因子 $\{Y_k\}_{k=1}^K \sim N(0, 1)$ 的线性组合. 即

$$Y = \sum_{k=1}^K w_k Y_k$$

其中

$$\sum_{k=1}^K w_k^2 = 1$$

定义

$$Z = \sum_{k=1}^K v_k Y_k$$

$\{v_k\}_{k=1}^K$ 只需满足

$$\sum_{k=1}^K v_k^2 = 1$$

由此, Z 仍然为标准正态分布, 将 $PL(Y)$ 和

$$\beta(Z) = \mathbb{E}[Recovery | D = 0, Z]$$

所受的影响因子分别考虑. 例如, 某个地区基准地

价的变动(可以视为系统风险因子), 可能会影响部分回收情况的回收率期望, 但不会影响整个资产组合低峰回收贷款的概率.

定理 2 对于同质资产组合, 定义整个不良贷款组合的总体回收率 $R^{(m)}$, 同式(4), 则

$$P(\lim_{m \rightarrow \infty} [R^{(m)} - \mathbb{E}(R^{(m)} | Y, Z)] = 0) = 1 \quad (8)$$

证明见附录.

1.3 低峰回收比率与系统风险因子的关系

利用经典的信用风险结构模型^④, 考察随机变量“低峰回收比率” D 与系统风险因子的关系.

令 V_{it} 为第 i 个公司在 t 时刻的资产价值, 则 V_{it} 满足随机微分方程

$$dV_{it} = \mu_i V_{it} dt + \sigma_i V_{it} dW_{it} \quad (9)$$

其中, $W_{it} \sim N(0, 1)$ 是标准布朗运动. 由随机微分方程, 易得

$$\log(V_{it}) = \log(V_{i0}) + \mu_i t - \frac{1}{2} \sigma_i^2 t + \sigma_i \sqrt{t} W_{it} \quad (10)$$

在 0 到 T 时刻, 给定阈值 K_i , 若 $V_{iT} < K_i$, 则回收率为 α (低峰回收), 若 $V_{iT} \geq K_i$, 则遵循正常的部分回收率过程, 服从高峰回收随机变量. 令

$$p_i = P[V_{iT} < K_i] = N(-c_i) \quad (11)$$

其中

$$c_i = \frac{\log V_{i0} - \log K_i + \mu_i T - \sigma_i^2 T/2}{\sigma_i \sqrt{T}}$$

若定义资产的对数收益率 r_i 为

$$r_i = \log(V_{iT}/V_{i0})$$

以上可以改写成, 在 0 到 T 时刻, 给定阈值

$$s_i = \log(K_i/V_{i0})$$

若 $r_i < s_i$, 落入图 3 水平线之下, 则回收率为 α (低峰回收), 若 $r_i \geq s_i$, 则遵循正常的部分回收率过程, 服从高峰回收随机变量 $\beta(Y)$. 由此, 可以对每笔不良贷款 i , 定义该笔违约贷款的回收统计量 R_i , R_i 服从 Bernoulli 变量 $B(1: p_i)$, 如式(1)所示, 其中 $p_i = PL_i$, 其分布如式(11)所给出. α 通常为接近于 0 的常数^⑤, $\alpha < \beta(Y)$, $\beta < \beta(Y)$ 依赖于系统风险 Y (宏观经济因素和行业地区等

④ 信用风险中结构模型, 参见 Merton^[7], 在 KMV 的模型也广泛使用.

⑤ 需要注意, 虽然有些贷款预计是回收率为零, 但是出于对业务人员“道德风险”的控制, 仍然有一个回收率要求, 所以 α 通常为接近于 0 的常数(如 1% ~ 5%).

系统风险).

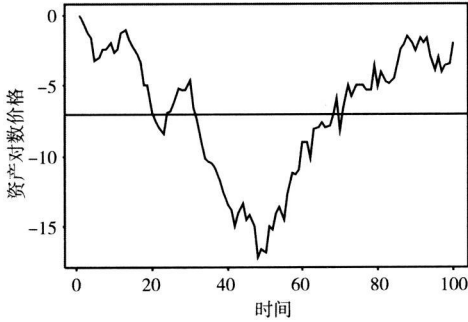


图 3 信用风险结构模型示例

Fig. 3 Example of credit risk structural model

定义第 i 笔违约贷款违约回收率为低峰 (即回收率为 α) 的示性变量为 D_i

$$D_i = \mathbb{I}_{P|V_{iT} < K_i} = \mathbb{I}_{P|r_i < s_i}$$

“低峰回收比率” D 如式 (2) 所定义. 由此, 通常选择 T 为 1 年, 假定第 i 笔违约贷款对应的公司在 1 年内资产收益率 r_i 为正态分布, 可以得到

$$s_i = N^{-1}(PL_i(Y))$$

对 r_i 建模, 假设其由单因子模型驱动, 其中 Y 为资产包中所有贷款都承担的系统风险, ε_i 为第 i 笔不良贷款的异质风险, 则 r_i 可以写成如下单因子模型形式:

$$r_i = \theta Y + \sqrt{1 - \theta^2} \varepsilon_i \quad 0 < \theta < 1 \quad (12)$$

其中, $Y \sim N(0, 1)$, $\varepsilon_i \sim N(0, 1)$, 则有任意一笔违约贷款 i 基于 Y 的条件低峰回收概率为

$$\begin{aligned} p(Y) &= PL_i(Y) = \mathbb{P}(D_i = 1 | Y) \\ &= \mathbb{P}(r_i < s_i | Y) \\ &= \mathbb{P}(\theta Y + \sqrt{1 - \theta^2} \varepsilon_i < s_i | Y) \\ &= N\left[\frac{s_i - \sqrt{\theta} Y}{\sqrt{1 - \theta^2}}\right] \end{aligned} \quad (13)$$

在巨额同质资产组合的假设下, 令 $s_i = s$, 对于每笔贷款, 其低峰回收概率均相等, 可以将式 (13) 改写为:

$$\begin{aligned} PL(Y) &= PL_1(Y) = PL_2(Y) = \dots \\ &= PL_n(Y) = N(c - eY) \end{aligned} \quad (14)$$

其中

$$\begin{aligned} c &= s / \sqrt{1 - \theta^2}, \\ e &= \sqrt{\theta} / \sqrt{1 - \theta^2} \end{aligned}$$

“低峰回收比率”是衡量不良贷款组合的重要指标. 在单因子模型的假设下, 类似于 Vasicek 的文献 [8] 和 [5] 的方法, 本文给出“低峰回收比率 D ”的分布函数和密度函数.

定理 3 定义违约贷款组合中“低峰回收比率” D , 如式 (2), $PL(Y)$ 为给定系统风险因子 Y 下的“低峰回收概率”, 则 D 的分布如下

$$\mathbb{P}[D \leq x | Y] = N\left(\frac{\sqrt{1 - \theta^2} N^{-1}(x) - N^{-1}(PL(Y))}{\theta}\right) \quad (14)$$

D 的密度函数为

$$f(x | Y) = \sqrt{\frac{1 - \theta^2}{\theta}} \exp\left(-\frac{1}{2\theta^2} \left(\sqrt{1 - \theta^2} N^{-1}(x) - N^{-1}(PL(Y))\right)^2 + \frac{1}{2} (N^{-1}(x))^2\right) \quad (16)$$

1.4 高峰回收率的预测方法和统计推断性质

上面阐述了高峰回收率

$$\beta(Z) = \mathbb{E}[Recovery | D = 0, Z]$$

($D = 0$ 表示基于所有非低峰回收情形的均值) 它是依赖于宏观经济周期等系统风险因子的随机变量. 如何估计“高峰回收率”和“低峰回收概率”之间的关系是本文需要解决的问题. 因为低峰回收概率可以用历史数据估计, 在通过有限数据样本估计出“低峰回收概率”, 则通过系统风险因子的共同作用, 预测出“高峰回收率”, 对回收工作具有重要的指导意义. 实际上, 通过部分处置不良贷款的“低峰回收概率”, 从而判断出在当年的“高峰回收率”, 继而可以预测当年未处置不良贷款的可能回收率推断, 从而达到对未处置不良贷款的回收率预测和风险管理.

在数据数量稀少的情况下, 希望采用尽可能简单的模型完成估计和预测, 模型如下

$$N^{-1}(\beta(Z)) = \alpha + \beta N^{-1}(PL(Y)) + \varepsilon \quad (17)$$

其中

$$\alpha = -a + \frac{bdc}{e}, \beta = -\frac{bd}{e}, \varepsilon = b \sqrt{1 - d^2} X \text{ 为}$$

残差.

证明 假设^⑥

$$\beta(Z) = 1 - N(a - bZ)$$

$$= N(bZ - a) \tag{18}$$

在 Y, Z 的条件期望下,可以得到如下形式

$$\begin{aligned} \mathbb{E}(R^{(m)} | Y, Z) &= PL(Y) \times \alpha + \\ &PH(Y) \times \beta(Z) \end{aligned} \tag{19}$$

显然可以对 Z 分解成独立随机变量和,如下

$$Z = dY + \sqrt{1 - d^2}X \tag{20}$$

其中

$$d = Cov(Y, Z)$$

由式(18)和式(20),可以得到

$$\begin{aligned} N^{-1}(\beta(Z)) &= bdY + \\ &b\sqrt{1 - d^2}X - a \end{aligned} \tag{21}$$

由式(14)

$$Y = \frac{c - N^{-1}(PZ)}{e}$$

代入式(21),最终得到式(17). 证毕.

显然,以上模型是一个简单的线形回归模型,参数估计是平凡的. $PL(Y)$ 可以表示为在不同宏观经济周期等因素下(系统风险因子 Y)的“低峰回收概率”,从而由样本的“低峰回收比率”作为估计. 而 $\mathbb{E}(Recovery | D = 0, Z)$ 可以作为每年的非极端回收的平均回收率,视为不同宏观经济周期等因素影响下的正常回收下的回收率均值. 依赖于模型式(17),在不同时间段,估计“低峰回收比率”,从而可以估计出该时间段的“高峰回收率”均值,用以估计未处置不良资产组合的回收率.

本文以 LOSSMETRICS 中 2000—2007 年的不良贷款数据为样本,考察正态逆变换前后的低峰回收概率和高峰回收率期望值的散点图. 可以看出,未变换前的图 4 中,低峰回收概率和高峰回收率期望值没有明显的趋势,而变换后的低峰回收概率和正常回收期望回收率可以用线形模型拟合. 图 5 中的直线为拟合的线性模型.

以上的实证结果表明,通过低峰回收概率估计高峰期望回收率的线形模型是合理有效的.

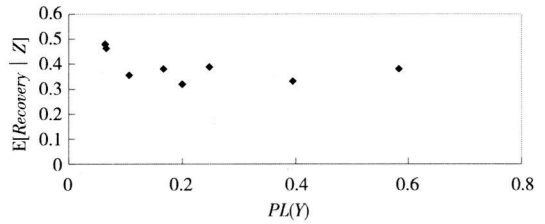


图 4 未变换低峰回收概率对高峰回收期望回收率散点图
Fig. 4 Scatter plot (low-recovery probability vs. expectation of high-recovery loans)—before transformation

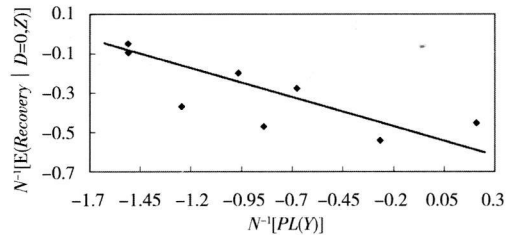


图 5 变换后低峰回收概率对高峰回收期望回收率散点图
Fig. 5 Scatter plot (low-recovery probability vs. expectation of high-recovery loans)—after transformation
给出模型的线性统计模型如下

Regression Statistics	
Multiple R	0.765 524 15
R Square	0.586 027 225
Adjusted R Square	0.517 031 762
Standard Error	0.418 649 199
Observations	8

	Coefficients	Standard Error	t Stat	p-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-1.632 6	0.306 0	-5.335 6	0.001 8	-2.381 4	-0.883 9	-2.381 4	-0.883 9
$N^{-1}[PL(Y)]$	-2.542 1	0.872 2	-2.914 4	0.026 8	-4.676 4	-0.407 8	-4.676 4	-0.407 8

模型具有统计意义上的显著性. 同时, 由于数据的特殊性质所限 (数据属于东方中和数据公司所有), 无法公开除模型以外的信息.

2 大规模同质不良贷款组合回收率分位数估计

针对不良贷款组合, 如何对不良贷款进行风险控制, 目前尚未有成形的体系和方法. 巴塞尔协议 II 中, 着眼于正常贷款组合的管理, 从而更关心违约概率和违约回收率的联合特征. 然后, 由于不良贷款组合中的违约概率均为 1 (按照 ISDA 的信用事件定义), 由此对于不良贷款组合的风险管理, 应该是基于回收率预测的风险管理, 其目的应着重在于监控“高峰回收率”和“低峰回收比率”的变化.

但是, 需要指出的是, 不良贷款组合的风险管理与通常意义下的风险管理有所不同. 实际上, 不良贷款组合回收率所关注的回收值分位数 (quantile) 的意义为: 给定某一分位点 (如 99%), 该组合处置结束后, 最终回收率高于该分位数值的可能性大于 99%. 其数学表述如下: 某不良贷款组合, 给定置信度 α , 可以定义不良贷款组合的 quantile 为

$$\text{Quantile}_\alpha(L) = \inf\{x \geq 0 \mid \mathbb{P}(L \leq x) \geq \alpha\} \quad (22)$$

其中 $L = \text{Cost} - R$, 即不良贷款组合的损失是商业化收购成本减去预计回收率.

在 1.3 节的模型下, 可以得到预计回收率的收敛解 R . R 由定理 4 给出. 定理 4 保证了当不良贷款组合足够大时, 其 quantile 值的可以直接由定理 4 中的公式和式 (17) 中的参数估计算出.

定理 4 根据不良贷款组合可以定义,

$$\begin{aligned} R &= N(c - eY) \times \alpha + \\ &\quad (1 - N(c - eY))N(bZ - a), \\ Z &= dY + \sqrt{1 - d^2}X, e > 0, bd > 0, \end{aligned}$$

则

$$\mathbb{P}(\lim_{m \rightarrow \infty} |R^{(m)} - R| = 0) = 1 \quad (23)$$

并且给出 R 的分布如下

$$\int_{B(x)} \int_{-\infty}^{A(x,y)} (\varphi(x) dx) \varphi(y) dy \quad (24)$$

其中, $\varphi(\cdot)$ 为标准正态分布密度函数.

$$A(x,y) = \frac{a - bdy - N^{-1}[(x - N(c - ey)) \times \alpha] / (1 - N(c - ey))}{b \sqrt{1 - d^2}}$$

$$B(x) = \frac{c - N^{-1}(x/\alpha)}{e}$$

特别的, 当 $\alpha = 0$ 时, R 的分布如下

$$\begin{aligned} \mathbb{P}[R \leq x] &= \\ 1 - \int_{B(x)} \int_{A(x,y)} (\varphi(x) dx) \varphi(y) dy &\quad (25) \end{aligned}$$

其中: $\varphi(\cdot)$ 为标准正态分布密度函数.

$$A(x,y) = \frac{a - bdy + N^{-1}[x/N(ey - c)]}{b \sqrt{1 - d^2}}$$

$$B(x) = \frac{N^{-1}(x) + c}{e}$$

为了估计分位数值, 还需要估计出 θ . 由式 (14) 可以得到 θ 的估计为

$$\theta = \sqrt{\frac{\text{Var}(N^{-1}(PL(Y)))}{1 + \text{Var}(N^{-1}(PL(Y)))}} \quad (26)$$

其中, $PL(Y)$ 可由样本“低峰回收比率”估计得出.

3 结 论

由于起步晚, 缺乏数据和经验的积累, 加之中国社会正处在全面转型之中, 存在着或多或少的结构性变化, 违约损失率及贷款组合的定价和风险管理可谓是任重而道远. 本文构建的针对不良贷款的巨额同质资产组合的分析框架, 包括对 Basel II 内部评级法中的重要信用风险要素——违约损失率 (LGD) 的计量方式、预测估计方法等问题仍然需要继续探讨. 就不良贷款组合风险分析的研究角度而言, 由于信用风险研究自身的复杂性和我国信用风险实证分析条件的不足, 使国内在该领域的研究进展相对缓慢, 尚没有形成广泛认可的理论体系和模型方法. 2007 年至今尚未平息的次贷危机, 也说明针对信用资产组合研究的必要性. 针对不良贷款的风险监控, 国内目前对操作风险、道德风险考虑得较多, 而对不良贷款的组合风险, 尚未有很好的研究. 本文从理论和实证角度, 对国内信用风险的主要载体——违约贷款的违约回收率建立理论框架, 并寻找实证研究的

突破口,将系统风险因子纳入建模框架,考虑时变系统风险因子,从而结合考虑违约损失率和宏观经济因素等深层次的联系;同时,从真实数据出发,建立低峰回收比率和整个不良贷款组合回收率之间的联系。

不良贷款组合的风险定价,目前于资产管理公司而言,在资产竞购、出售配置和风险控制等方面,具有重大的参考意义和现实价值。对于银行的信贷资产组合管理,也提供了可以比对的方法和对象。对于今后的信贷资产证券化、信用衍生

品——如基于不良贷款资产池的抵押担保证券(CDO)、信贷资产的结构性和信托分层产品等创新业务,初步提供了其定价的理论基础和数学框架。

由于中国银行业信贷业务的特殊性,不能盲目地照搬国外的建模方法,应当在借鉴国外各种优秀方法的基础上,结合中国本地违约损失率的数据特征,探讨合适的模型构造基础、估计方法和检验手段,建立符合中国不良资产管理实际的信用风险模型和风险监控体系。

参考文献:

- [1] Hu Y-T, W Perraudin. The Dependence of Recovery Rates and Defaults. Birkbeck College and Bank of England[M]. Working Paper, 2002
- [2] 王春峰, 万海晖, 张维. 边界 Logistic 违约率模型及实证研究[J]. 管理科学学报, 1998, 1(1): 68-72
Wang Chun-feng, Wan Haihui, Zhang Wei. Credit risk assessment in commercial banks and its test[J]. Journal of Management of Sciences in China, 1998, 1(1): 68-72 (in Chinese)
- [3] 梁世栋, 郭冰, 方兆本. 随机违约强度下的信用风险期限结构研究[J]. 管理科学学报, 2005, 8(4): 75-79
Liang Shidong, Guo Bing, Fang Zhaoben. Study of credit risk term structure with stochastic default intensity[J]. Journal of Management of Sciences in China, 2005, 8(4): 75-79 (in Chinese)
- [4] 石晓军, 任若恩, 肖远文. 边界 Logistic 违约率模型及实证研究[J]. 管理科学学报, 2007, 10(3): 45-51
Shi Xiaojun, Ren Ruoen, Xiao Yuanwen. Bounded logistic default risk model and empirical evidences from China[J]. Journal of Management of Sciences in China, 2007, 10(3): 45-51 (in Chinese)
- [5] Vasicek O. The distribution of loan portfolio value[J]. RISK, 2002, 15(12): 128-146
- [6] Martin H. Modeling and estimating dependent loss given default[J]. RISK, 2006 (9): 323-339
- [7] Merton R C. On the pricing of corporate debt: The risk structure of interest rates[J]. Journal of Finance, 1974 (29): 449-470
- [8] Vasicek O. Limiting Loan Loss Probability Distribution[M]. KMV Press KMV Corporation, 1991
- [9] Bluhm C, Overbeck L, Wagner C. An Introduction to Credit Risk Modeling[M]. Chapman & Hall/CRC, 2003
- [10] Demine J, Neto de Carvalho C. Bank loan losses given default: A case study[J]. Journal of Banking & Finance, 2006, 30(4): 1219-1243
- [11] Gabriel J, Javier M. Modelling the distribution of credit losses with observable and latent factors[J]. Journal of Empirical Finance, 2009, 16(2): 235-253
- [12] Darrell D, Kenneth J Singleton. Credit Risk: Pricing, Measurement, and Management[M]. Princeton Press, 2003
- [13] Gupton G, Stein R. LossCalc: Moody's Model for Predicting Loss Given Default (LGD)[M]. Moody's Investor Service, 2002
- [14] Jones E, Philip S P. Mason and Eric Rosenfeld. Contingent claims analysis of corporate capital structures: An empirical investigation[J]. Journal of Finance, 1984, 39(3): 611-625
- [15] Gupton G, Gupton C. Bank loan loss given default[J]. Moody's Investor's Service Global Credit Research, 2000 (11): 1-24
- [16] William H G. Econometric Analysis[M]. Fifth Edition, Prentice-Hall, 2002
- [17] 中科院, 东方中和数据公司. 东方中和公司不良资产回收情况研究报告[M]. 2007
Chinese Academy of Science, Zhonghe Data Consult Company. Report of Orient AMC non-performance asset[M]. 2007 (in Chinese)

[18] 刘志刚, 罗光, 杨志明. 违约损失率数据库与预测模型的构建[J]. 金融研究, 2007, (05B): 33-42

Liu Zhigang, Luo Guang, Yang Zhiming. LGD database and prediction model[J]. Journal of Financial Research, 2007, (05B): 33-42 (in Chinese)

Approach to estimate loss given default of large homogenous non-performing loan portfolio

MA Yu-chao, CHEN Mu-zi, CHEN Hao, WANG Bao, CHEN Min, YANG Xiao-guang

Chinese Academy of Science, Academy of Mathematics and System Science, Beijing 100190, China

Abstract Recovering non-performing loan is a key issue in the innovation and reform of China financial industry. Batch-package method is one of the first alternatives to do with non-performing loan recovering. The core issue in the recovering procedure is to estimate Loss Given Default (LGD) of non-performing loan portfolios. Based on the biphasic feature of recovery rate, we classify non-performing loans into low-recovery loans and high-recovery loans, and prove that the recovery rate of non-performing loan converges to the sum of low-recovery rate and conditional expectation of high-recovery rate. Using classical structural model, we further derive that there is linear relation between the inverse-normal transformation of conditional expectation of high-recovery rate and the inverse-normal transformations of low-recovery rate. Since it is easy to discriminate the low recovery loan, and hence easy to estimate low-recovery rate, we can estimate the whole non-performing portfolio recovery rate by linear regression. Then we give an approach to estimate the Loss Given Default of large homogenous non-performing loan portfolio. Based on the model, we give a method to calculate the quantile of non-performing loan portfolio recovery rate, which is an extension of VaR.

Key words non-performing loan, large homogenous portfolio, recovery rate estimation, structural model

附录

1 定理 1 的证明:

证明 该思想来源于 (Bluhm^[9]).

为简化符号, 令

$$\eta_i = \alpha p_i + \beta(1 - p_i)$$

固定, $y \in \mathbb{R}$, 定义条件概率

$$P_y: P_y = P[\cdot | Y = y]$$

令

$$X_k = N_k[\eta_k R_k - \mathbb{E}(\eta_k R_k | Y)]$$

则 $\{X_k\}_{k=1}^m$ 互相独立. 令

$$\tau_m = \sum_{i=1}^m N_i$$

显然, $\{\tau_m\}_{m=1}^{\infty}$ 为严格递增序列. 由强大数律, 只需满足条件

$$\sum_{k=1}^{\infty} \frac{1}{\tau_k^2} \mathbb{E}[X_k^2] < \infty \quad (27)$$

即可证明. ©2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$$P_y[\lim_{m \rightarrow \infty} \frac{1}{\tau_m} \sum_{k=1}^m X_k = 0] = 1 \quad (28)$$

式(28)对任意 $y \in \mathbb{R}$ 成立, 由 a.s 收敛和依概率收敛关系, 式(7)显然成立.

以下证明式(27)成立. 注意到由定义 $\eta_k < 1$, 且 $R_k < 1$, 有

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{\tau_k^2} \mathbb{E}[X_k^2] &= \sum_{k=1}^{\infty} \frac{1}{\tau_k^2} \mathbb{E}[N_k^2[\eta_k R_k - \mathbb{E}(\eta_k R_k | Y)]^2] \\ &\leq \sum_{k=1}^{\infty} \frac{1}{\tau_k^2} N_k^2 \mathbb{E}[\eta_k R_k - \mathbb{E}(\eta_k R_k | Y)]^2 \\ &\leq \sum_{k=1}^{\infty} \frac{1}{\tau_k^2} 4 \times N_k^2 < \infty \end{aligned} \quad (29)$$

证毕.

2 定理 2 的证明: 类似于定理 1, 此处略去.

3 定理 3 的证明:

证明

$$P[D \leq x] = P[PL(Y) \leq x]$$

$$= P[Y \geq PL^{-1}(x)]$$

$$= N(-PL^{-1}(x)) \tag{30}$$

将式(14)代入,可以得到式(15). 对

$$f(x) = \frac{dP(D \leq x)}{dx}$$

求导,易得式(16).

4 定理4的证明:

证明 直接由定理1和定理2可得出式(23).

由 $R^{(m)}$ 的定义

$$P[R \leq x]$$

$$\begin{aligned} &= P[N(c - eY) \times \alpha + (1 - N(c - eY))N(-a + bY + b\sqrt{1 - d^2}X) \leq x] \\ &= E[1_{|N(c - eY) \times \alpha \leq x|} \times 1_{|N(-a + bY + b\sqrt{1 - d^2}X) \leq (x - N(c - eY) \times \alpha) / (1 - N(c - eY))|}] \\ &= E[1_{|Y \geq [c - N^{-1}(x/\alpha)]/e|} \times 1_{|N(-a + bY + b\sqrt{1 - d^2}X) \leq (x - N(c - eY) \times \alpha) / (1 - N(c - eY))|}] \\ &= E[1_{\{Y \geq \frac{c - N^{-1}(x/\alpha)}{e}\}} \times 1_{\{X \leq \frac{a - bY - N^{-1}[(x - N(c - eY) \times \alpha) / (1 - N(c - eY))]}{b\sqrt{1 - d^2}}\}}] \\ &= \int_{B(x)} \int_{-\infty}^{A(x,y)} (\varphi(x) dx) \varphi(y) dy \end{aligned} \tag{33}$$

其中

$$\begin{aligned} A(x,y) &= \frac{a - bdy - N^{-1}[(x - N(c - ey) \times \alpha) / (1 - N(c - ey))]}{b\sqrt{1 - d^2}} \\ B(x) &= \frac{c - N^{-1}(x/\alpha)}{e} \end{aligned} \tag{34}$$

当 $\alpha = 0$ 时

$$R = N(eY - c)N(bZ - a)$$

$$P[R \leq x]$$

$$\begin{aligned} &= P[(1 - N(c - eY))N(-a + bY + b\sqrt{1 - d^2}X) \leq x] \\ &= P[N(eY - c)N(-a + bY + b\sqrt{1 - d^2}X) \leq x] \\ &= E[1_{|N(eY - c) \leq x|} + 1_{|N(eY - c) > x|} \times 1_{|N(eY - c)N(-a + bY + b\sqrt{1 - d^2}X) \leq x|}] \\ &= 1 - E[1_{|N(eY - c) > x|} \times (1 - 1_{|N(eY - c)N(-a + bY + b\sqrt{1 - d^2}X) \leq x|})] \\ &= 1 - E[1_{|N(eY - c) > x|} \times (1_{|N(eY - c)N(-a + bY + b\sqrt{1 - d^2}X) > x|})] \\ &= 1 - E[1_{|Y > \frac{N^{-1}(x) + c}{e}|} \times (1_{\{X > \frac{a - bY + N^{-1}[x/N(eY - c)]}{b\sqrt{1 - d^2}}\}})] \\ &= 1 - \int_{B(x)} \int_{A(x,y)}^{+\infty} (\varphi(x) dx) \varphi(y) dy \end{aligned} \tag{35}$$

其中

$$\begin{aligned} A(x,y) &= \frac{a - bdy + N^{-1}[x/N(eY - c)]}{b\sqrt{1 - d^2}} \\ B(x) &= \frac{N^{-1}(x) + c}{e} \end{aligned} \tag{36}$$

证毕.