

外部数据度量中国商业银行操作风险的样本量研究

高丽君¹,宋汉鲲^{2,3}

(1. 山东财经大学 工商管理学院,山东 济南 250014;2. 青岛大学 商学院,山东 青岛 266071;
3. 中国工商银行青岛市分行 山东路支行,山东 青岛 266000)

摘要:随着巴塞尔新资本协议 II、III 的陆续出台,中国银监会也要求国内银行必须对操作风险计提监管资本。但计提监管资本需要准确度量风险,而中国学者因难以获得内部数据只好采用外部数据进行度量,每个研究团队搜集的外部数据是否充分?多少外部数据能够体现总体特征?最少需收集多少外部数据?至今尚无人回答这些问题。文章就此问题采用非参数变规模拔靴法,通过设定各分位数均值变化率截点值来选择样本规模,对每一样本规模多次模拟后拟合损失强度分布并模拟年度操作风险资本金,经过对比分析后,得出最小样本规模在 800 例(包含极端值),模拟结果比较稳定。与 Wind 数据库对中国银行业 2013 年的操作风险资本金判断相比,度量结果比较合理。这说明,不考虑所收集的外部数据是否能代表总体特征就计算,结果可能会产生一定偏差。

关键词:商业银行;操作风险;外部数据;样本量;变规模拔靴法

中图分类号:F832.332 **文献标识码:**A **文章编号:**2095-929X(2015)03-0068-08

一、问题提出

操作风险是银行固有的重要风险之一。自巴塞尔新资本协议 II、III 系列文件的陆续出台,国际学术界及业界迎来了两次操作风险研究的热潮。中国银行业操作风险大案要案频发,给银行界带来极不利影响,中国银监会(简称银监会)要求国内银行自 2013 年起必须计提操作风险资本金^[1]。诸多因素表明,操作风险研究具有现实紧迫性,理论界和银行界具有开发操作风险度量模型的动力,操作风险已经成为风险管理中一个非常重要的领域。

对操作风险计提资本金是操作风险管理的第一支柱的内容,是操作风险管理的最后一道防线。资本度量需要充足的、切实可信的数据作基础。然而中国商业银行与国际银行业在度量银行操作风险上的一个最重要的不同在于:尽管国际银行业的操作风险数据也是不足的,但由于数据收集工作入手比较早,有一定的可信的

收稿日期:2014-11-11

基金项目:国家自然科学基金项目“外部数据下中国商业银行操作风险度量模型验证研究”(71301087),山东省自然科学基金项目“基于损失属性的中国商业银行操作风险管理研究”(ZR2010GL011)。

作者简介:高丽君,女,湖北武汉人,博士,山东财经大学工商管理学院教授,研究方向:风险管理;宋汉鲲,男,山东济南人,青岛大学商学院硕士生,中国工商银行青岛市分行山东路支行,研究方向:战略管理。

内部数据基础,有相应的国际监管机构按一定标准收集的、行业数据库或企业数据库的外部数据作为补充,研

表 1 国内学者估算中国商业银行操作风险资本金(亿元)的研究结果对比

方法隶属框架	计量方法	样本区间	样本数	估计资本金	置信度	作者	发表年度	发表刊物	度量结果验证
损失分布法(LDA)	LDA + 模拟	1990 – 2003	71	1900	99.9%	樊欣、杨晓光	2005	系统工程理论与实践	—
极值理论	HKKP 估计	1994 – 2005	193	13624	99.9%	高丽君、李建平、徐伟宣等	2006	系统工程	占银行业总资产比
极值理论	POT 极值	1994 – 2005	193	13089	99.9%	高丽君、李建平、徐伟宣等	2007	运筹与管理	占银行业总资产比
模拟	蒙特卡罗模拟	2000 – 2006	578	3163	99.9%	费伦苏、邓明然	2007	金融论坛	—
模拟	灰色动态残差模拟	1994 – 2005	219	1733.87	99.9%	金婷、秦学志	2007	价值工程	—
LDA	LDA	1990 – 2005	365	107	99.9%	张宏毅、陆静	2008	系统工程学报	—
极值、贝叶斯	贝叶斯推断法	2003 – 2008	174	697.9	99.9%	卢安文、任玉琰、唐浩阳	2009	系统工程学报	—
极值理论	POT 幂率模型	1995 – 2006	868	9700	99.9%	司马则茜、蔡晨、李建平	2009	中国管理科学	与 HKKP 法比误差
极值理论	BMM 模型	1994 – 2008	220	31.91	99.9%	钱艺平、林祥、陈治亚	2010	统计与决策	—
极值理论	POT 极值	1995 – 2006	860	1236	99.9%	丰吉闯、李建平、高丽君	2011	国际金融研究	与 GED、SGED 比
LDA	GED	1995 – 2006	860	1201	99.9%	丰吉闯、李建平、高丽君	2011	国际金融研究	与 POT、SGED 比
LDA	SGED	1995 – 2006	860	1057	99.9%	丰吉闯、李建平、高丽君	2011	国际金融研究	与 POT、GED 比
LDA	LDA + 估计阈值	2000 – 2006	280	1930	99.9%	丰吉闯、李建平、陈建明	2011	管理评论	压力测试
LDA	LDA	1994 – 2009	409	39	99%	周艳菊、彭俊、王宗润	2011	中国管理科学	与 MLE 法比较
LDA	LDA + 相关性	1994 – 2009	409	32	99%	周艳菊、彭俊、王宗润	2011	中国管理科学	与不相关法比较
极值理论	变点分析阈值法	1987 – 2010	343	749.4	99.9%	宋坤、宋鹏	2011	统计与决策	—
极值理论	g – h 分布	1995 – 2006	868	8780	99.9%	司马则茜、蔡晨、李建平	2011	系统工程理论与实践	与 POT 幂率法比
极值理论	POT + Copula	1994 – 2007	202	4.99	99.9%	吴恒煜、赵平、严武等	2011	运筹与管理	与不相关法比
极值 + 模拟	贝叶斯参数估计	1987 – 2011	344	1080.1	99%	宋坤、刘天伦	2012	统计与信息论坛	与 MLE 法比标准差
极值理论	BMM	1990 – 2009	439	11.95	99.9%	陆静	2012	管理工程学报	KS 及 AD 检验
极值 + LDA + 模拟	BS – PSD – LDA 模型	1994 – 2010	426	16.91	99%	王宗润、汪武超、陈晓红等	2012	管理科学学报	Kupiec 返回检验
极值 + 相关性	极值 + 多元 COPULA	1990 – 2010	238	1097.9	99.9%	陆静、张佳	2013	中国管理科学	与简单极值、LDA 比
极值 + 相关性	尾相关 COPULA	不明	200	883.22	99.9	明瑞星、谢铨	2013	统计与决策	与不考虑尾相关比

注:表格数据根据各研究文献统计整理。

研究者关注的重点在于内、外部数据的合并问题^[2-3]。而中国的商业银行,内部操作损失数据收集起步晚,甚至有的银行还未提上议事日程^[4],未开发出相应的行业数据库,银行监管机构也未系统进行操作风险损失数据的收集(仅银监会内部有少量关于案件稽核的信息但不公开)。这导致对中国商业银行操作风险度量的研究,绝大多数学者采用外部数据,将整个中国商业银行看作一家大银行进行度量与分析(见表1)^[5-24]。而对于究竟需为中国商业银行分配多少资本金这个问题,中国学者众说纷纭没有统一的答案,最小者5亿元,最高者1.3万亿元,导致对中国商业银行操作风险进行度量的模型说服力不强,业界也难以采用研究者所提出的模型。究其原因,主要有以下几个方面:第一,操作风险内涵广泛,巴塞尔委员会并未提出内涵非常明晰的定义,不同研究者对操作风险的理解不同,导致收集的操作损失案例不同,数据标准亦不同;第二,由于我国商业银行内部数据应保密,不易收集,多数研究采用外部数据直接度量,每个团队收集的外部数据都难以穷尽所有,均是小样本的,是否能够代表总体样本情况,并不清楚;第三,外部数据本身具有偏差,而目前国内研究者均是将外部数据直接纳入模型的,不修正外部数据的内生偏差,度量操作风险会产生内外数据不同质、扩大规模问题和缺乏兼容性等一系列问题,可能会对结果造成重大影响;第四,采用的度量模型不同,模型能否较准确度量操作风险,近半数未验证,进行验证的多数为与其它模型互比结果,验证方法单一。

以上说明中国商业银行操作风险的度量存在一系列未决问题,即至少存在三个问题:一是如何保障或证明,现有的外部数据能够代表总体特征?二是外部数据的修正问题;三是如何选择或开发有效的模型,能从这些非常有限的数据中,发现与实际接近的操作风险值,并验证模型。这表明操作风险度量,数据本身的研究非常重要,数据是操作风险度量与管理的基础性工作,需要收集具备何种基本特征的外部数据,至少收集多少才能体现总体特征的?外部数据内生偏差如何处理?需要深入探讨。在数据规范、充分、准确基础上,模型的适用性、准确性、一致性亦需要验证,以加强模型说服力,促使业界采纳。

二、研究思路

该文拟立足第一个问题,即在外部数据小样本情况下,研究其是否能够代表总体样本特征。即,至少需收集多少有效的外部数据(包含极端值),才能体现总体特征?

(一)前提假设

该研究在以下几个假设条件下进行分析:

假设1:中国研究者们经过多年对巴塞尔新资本协议的理解和对操作风险的研究,对操作风险的范畴的理解及掌握情况基本达到了一致。即,可能各研究团队所收集到的操作风险外部数据案例不同,但假设已经收集到的操作损失外部数据数据标准基本保持一致。

假设2:在中国社会背景条件下,可以认同各商业银行的操作运营具有一定的同质性,可以将各商业银行看作一个大的中国商业银行。

假设3:过去发生的损失今后仍有可能发生,且损失额度与发生年份之间的关系不大(高丽君,郭艳丽利用内部欺诈数据证明了损失额度与损失发生年代之间的影响较小^[25])。在此假设下,可以不考虑过去发生的损失额度的时间效应。

假设4:外部数据确有内生偏差,但中国学者收集到的外部数据多为公共数据(因为中国无相关行业数据库,也很少有相应的保险数据),在假设2的前提下,通过虚拟一家大的中国商业银行可以不考虑控制偏差和尺度偏差。即,只需考虑报告偏差。报告偏差可以通过某些修正技术进行数据预处理加以修正。

假设5:银行可以采用外部损失数据来测算操作风险所需的资本要求。Nash指出,有证据显示如果商业银行不采用外部损失数据来弥补数据不足的话,就不能准确地估计出操作风险分布的尾部^[26],这很难让监管

当局相信度量方法的可行性,所以商业银行采用外部损失数据是不可避免的。Fontoouvelle 等^[27]采用 OpRisk Analytics 和 Op Vantage 两家数据公司从公共信息渠道收集的损失数据,对国际活跃银行所需提取的操作风险资本金进行了模拟测算,结果表明利用外部数据能很好的测算操作风险所需的资本要求。所以,出于数据缺乏的现状,量化操作风险时需要合并单个金融机构以外的外部数据。

假设 6:通过随机抽样技术多次抽取样本,并对各在险值取均值,即使是小样本,也能抽取到极端值。
在此假设前提下,思考需要收集的外部数据量,即,至少需要收集多少有效的外部数据样本,才能体现总体特征。

(二)研究方法

采用非参数变规模拔靴法生成不同的模拟样本量,考虑各分位数、尾部在险值的变化情况,选择能反映总体特征的比较稳定的最小样本量。

Haubenstock 和 Hardin 认为当损失事件个数大于 100,并且包括损失非常大的数据(尾部)时,就可以直接对风险建模,使用的数据量越大,度量结果越稳定^[28]。Embrechts 认为数据量的不足会对数据的平稳性以及结果产生严重不利影响^[29]。因此,本文首先需要考虑备选初始样本规模,初始样本规模从 100 开始。其次,需要考虑单次规模增加量,这里,本文选取 100。

利用样本数据集,以非参数 bootstrap 法变规模分析,探求最小样本数多大时,总体样本比较稳定? 具体思路:样本数据集有放回抽样,从中抽取 100 – 10000 个样本数,每个规模抽取一定次数(本文抽取 5000 次),分析每个规模的新样本的特征。由于操作风险具有显著的尖峰厚尾特征,本研究分析的重点显然不应如正态分布那样分析均值、中位数、方差等。本研究分析分布的总体特征,考虑四分位数、尾部极端值的变化情况。当各分位数比较稳定时,可以认为该样本量基本能体现总体特征。选取达到稳定的最小规模量作为所求最小的能体现总体特征的样本量。这里,本文视均值变化率来表征稳定,当某样本规模的所有在险值的均值的变化率小于一定值时,认为达到稳定。这个取样规模即所求的能代表总体特征的最小样本规模。

三、实证分析

(一)原始外部样本数据描述

研究组共从外部公共媒体如报纸、财经网络、中国银监会网站、审计署、法院案例收集到从 1994 – 2013 年的数据超 3000 例。考虑到收集成本,低于一万元人民币的损失暂不列入数据库,即,本数据库为左截断数据。鉴于 2013 年的数据,由于某些事件仍未侦查完毕,损失额度可能不准确,暂不考虑。另有部分事件存在信息不全、最终损失为零(因发现及时或司法介入最终收缴等原因导致银行最终未损失)或最终损失不明(仍在持续收缴)等情况,也暂不列入。最终纳入度量的原始样本为 1994 – 2012 年的单件损失共 2323 例。由表 2 可知,样本数据损失跨度非常大,平均单件损失超亿元,损失分布存在严重的尖峰厚尾现象。

表 2 外部数据样本集描述(单位:万元)

样本数	最小值	最大值	均值	中位数	标准差	偏度	峰度
2323	1	800000	10586	192.5	48070.5	8.878	96.669

注:其中 2009 年以前的数据见 www.casipm.ac.cn。相关数据收集标准及规范见 <http://www.cord-oprisk.com>。

(二)确定最小样本规模

表 3 是利用非参数 bootstrap 法变规模分析,将原始样本抽样 5000 次得出的分布分位数,限于篇幅,表中仅列了其中几个样本规模的情况。

为确定最小的能体现总体样本特征的样本量,考虑各在险值均值最大变化率。其中:最大均值变化率 = $\max\{\text{变化率}(\text{Var05}, \text{Var25}, \text{Var50}, \text{Var75}, \text{Var90}, \text{Var95}, \text{Var99}, \text{Var99.9}, \text{Var99.99})\}$ 。这里选取最大均值首次出现小于 0.001 的规模作为能体现总体样本的最小规模。当取样数达到 800 时,首次出现最大均值变化率小于 0.001,为 0.0009。表 4 是各取样规模下在险值最大均值变化率情况,图 1 是各样本规模不同在险值均值的变化率。因此初步确定的最小取样规模为 800 例。

表 3 原始样本非参数抽样模拟 5000 次的在险值(从 100 至 10000,单位:万元)									
取样数		var25	var50	var75	var90	var95	var99	var99.9	var99.99
100	均值	22.0901	209.8465	2733.0911	16921.2116	42768.07974	184149.9505	325570.2256	339712.3
	方差	96.2454	12954.35	2111958.853	62824778.25	610690048.8	16340352575	33756759021	3.75E + 10
200	均值	20.49306	194.0766	2576.07265	16491.31425	43092.22084	205558.2947	416371.0788	444687.4
	方差	36.6512	4680.795	814618.7984	26006419.6	292990765	12012215050	27955600705	3.31E + 10
.....									
5000	均值	19.38378	188.308	2481.163172	16179.75371	43458.57376	227151.1846	597974.9236	19.38378
	方差	1.487044	153.3175	31136.07685	1104420.592	10200785.47	999819106.4	3987182769	2.7E + 09
.....									
10000	均值	19.44623	189.0743	2477.119988	16155.39036	43430.56709	225092.5303	602342.5745	789390.3
	方差	0.719219	107.0605	12935.33713	648890.294	5267462.73	559359697.4	2323977055	1.25E + 09

表 4 各样本规模下在险值最大均值变化率

取样规模	100	200	300	400	500	600	700	800	900	1000
最大均值变化率	——	0.0751	0.0150	0.0098	0.0094	0.0053	0.0095	0.0009	0.0259	0.0098
取样规模	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000
最大均值变化率	0.0072	0.0009	0.0023	0.0114	0.0032	0.0046	0.0059	0.0023	0.0028	0.0044
取样规模	2100	2200	2300	2400	2500	2600	2700	2800	2900	3000
最大均值变化率	0.0052	0.0048	0.0010	0.0000	0.0060	0.0005	0.0028	0.0046	-0.0001	0.0039
取样规模	3100	3200	3300	3400	3500	3600	3700	3800	3900	4000
最大均值变化率	0.0043	-0.0012	0.0048	0.0028	0.0020	0.0012	0.0015	0.0020	0.0080	0.0003
取样规模	4100	4200	4300	4400	4500	4600	4700	4800	4900	5000
最大均值变化率	0.0054	0.0019	0.0021	0.0015	0.0030	0.0021	0.0004	0.0010	0.0026	0.0047

注:限于篇幅,样本规模大于 5000 的此处未列。

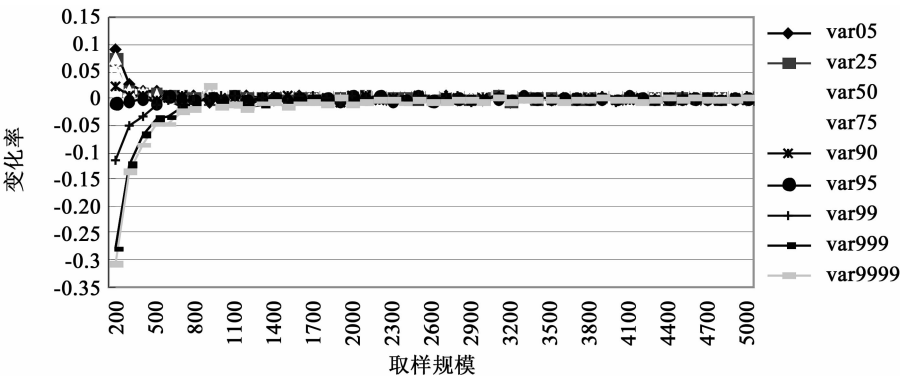


图 1 取样规模与各在险值均值变化率图

(三) 利用样本数据集的数据分析损失频率

国际上研究操作风险损失频率一般采用泊松分布或负二项分布。损失频率分布根据 Kolmogorov -

Smirnov(K - S) 检验选择分布形式(负二项分布), 模拟损失频率分布。表 5 是原始样本数据损失频率分布估计。根据估计的损失频率分布模拟生成损失频率随机分布。

分析样本数据集的损失频率。表 5 是由原始数据样本集每年发生的损失数判断损失频率服从分布。模拟服从负二项分布的损失频率发生的随机数 10000 个, 得出 10000 个单位时间发生的损失次数(这里, 本文视单位时间为 1 年)。

(四) 根据估计出来的最小样本量模拟损失强度分布

估计损失强度分布。采用非参数 bootstrap 方法有放回抽样 800 个, 模拟 10000 次获得 10000 组模拟数据。对每一组数据进行排序, 使得 $t[i-1,j] \leq t[i,j] \leq t[i+1,j], i=2, \cdots, 799, j=1, \cdots, 10000$ 。再对每一组已排好序数据按顺序取平均值, $t[i] = \frac{1}{10000} \sum_{j=1}^{10000} t[t,j]$ 。获得模拟的平均损失强度。表 6 为模拟的平均损失强度的分布情况。

(五) 联合分布估计及进一步分析

利用蒙特卡罗模拟得出损失联合分布。得到损失分布如表 7 所示的分位数及在险值 99.9 的均值。

表 7 模拟联合损失分布在险值均值(单位: 万元)

取样规模	Var05	Var25	Var50	Var75	Var95	Var99	Var99.9	Var99.99
400	167502	652578	1397886	3009735	9858056	33525557	100331295	1318285301
800	100547	487986	997926	1966726	6236727	17555229	57485541	170847407
2400	125961	439691	915957	1835201	5290181	12367790	56213740	129869540

考虑不同样本规模对最终结果的影响, 本文再分析不同预设最大均值变化率截点下, 得出的各取样规模(结合表 4) 及根据实证分析中(三)(四) 部分中的方法(仅取样规模不同) 模拟的损失强度分布。本例中, 各规模下损失强度分布通过 K - S 检验均可认为服从对数正态分布(其他常见分布形式如威布尔分布、正态分布、泊松分布等均未通过拟合优度检验)。表 8 为选择不同最大均值变化率截点所确定的取样规模, 及由此获得的平均损失强度的分布估计。

由表 7 及表 8 可以看出, 当取样规模达到 800 例后, 计算结果相对比较稳定, 考虑到样本收集的难度及各种成本(时间、人工、财物力等成本) 问题, 确定最小外部数据取样规模为 800 例且其中包含极端损失时, 可以体现总体特征。

根据模拟结果, 按照巴塞尔新资本协议 Var99.9% 的要求, 中国商业银行应计提操作风险资本金 5748.55 亿元人民币。根据 Wind 数据库对中国 2013 年银行业的风险加权资产(四个季度), 按照巴塞尔新资本协议计提操作风险资本不应低于风险加权资产 10.5% 的要求, 操作风险的资本要求大小约在 5197 - 6208 亿之

表 5 原始样本集损失频率分布估计

分布	参数估计值		D	P value
泊松分布	122.2105		0.366	0.0123
负二项分布	3.2996	122.2105	0.2418	0.2164

表 6 取样规模为 800 例时平均损失强度分布

分布估计	对数正态		威布尔	
	5.2128	3.0064	0.3178	936.53
拟合检验	0.15		< 0.05	

表 8 不同最大均值变化率截点所确定的取样规模及损失强度分布估计

最大均值变化率截点	取样规模	分布	损失强度分布估计	
			参数 1	参数 2
0.01	400	对数正态	5.5027	3.0754
0.005	800	对数正态	5.2128	3.0064
0.001	800	对数正态	5.2128	3.0064
0.0005	2400	对数正态	5.4077	2.9063
0.0002	2400	对数正态	5.4077	2.9063
0.0001	2400	对数正态	5.4077	2.9063

间,而按照最小取样规模为 800 例估计的操作风险资本金为 5749 亿元,因此可以认为度量比较准确。即,首先分析外部数据是否能够体现总体特征来确定最小样本规模再模拟得出的操作风险资本金相比直接对一组外部数据进行拟合分布估计,度量的资本金更准确(与表 1 对比)。此结果亦表明,此方法估计的能代表总体特征的最小样本数比较准确。

四、结 论

该文章针对“采用外部数据分析中国商业银行的操作风险,至少需收集多少数据才能够体现总体特征?”这个问题进行分析,采用非参数变规模拔靴法,设定在险值均值变化率截点值,根据最早出现的小于截点值的最大在险值均值变化率来选择样本规模,模拟一定次数后利用新生成的样本的排序均值拟合损失强度分布并模拟年度操作风险资本金,经过比较分析后,得出最小样本规模在 800 例(包含极端值),模拟结果比较稳定。与 Wind 数据库对中国银行业的操作风险资本金判断相比,度量结果比较合理,说明外部数据的取样规模对中国商业银行操作风险资本金度量具有一定的影响,不考虑所收集的数据是否能体现总体特征就进行计算,会产生一定的偏差。

参考文献:

[1]中国银行业监督管理委员会(简称中国银监会). 中国银行业实施新资本协议指导意见[R]. http://www.cbrc.gov.cn/gov-View_C0015CF9E28A41198373752F0C8631BA.html,2007-2-28.

[2]HELA D, GEORGES D. Scaling Models for the Severity and Frequency of External Operational Loss Data [J]. Bank & finance. 2010(34):1484-1496.

[3]JIM G, MONSERRAT G, JENS P N, et al. Using External Data in Operational Risk [J]. The Geneva Papers, 2007(32):178-189.

[4]曲绍强. 关于我国操作风险损失数据库构建的思考[J]. 商业研究,2006(17):149-151.

[5]樊欣, 杨晓光. 我国银行业操作风险的蒙特卡罗模拟估计[J]. 系统工程理论与实践,2005, 22(5): 44-48.

[6]费伦苏,邓明然. 商业银行操作风险的统计特征及其资本模拟实证[J]. 金融论坛,2007(8):3-11.

[7]丰吉闯, 李建平, 陈建明. 基于左截尾数据的损失分布法度量操作风险:以中国商业银行为例[J]. 管理评论, 2011,23(7):171-176.

[8]丰吉闯, 李建平, 高丽君. 商业银行操作风险度量模型选择研究[J]. 国际金融研究, 2011(8):88-96.

[9]高丽君, 李建平, 徐伟宣, 等. 基于 POT 方法的商业银行操作风险极端值估计[J]. 运筹与管理,2007 (1):112-117.

[10]高丽君, 李建平, 徐伟宣, 等. 基于 HKKP 估计的商业银行操作风险估计[J]. 系统工程, 2006, 24(6): 58-63.

[11]金婷, 秦学志. 用改进的蒙特卡罗模拟估计我国商业银行操作风险[J]. 价值工程, 2007(12):10-14.

[12]卢安文,任玉珑,唐浩阳. 基于贝叶斯推断的操作风险度量模型研究[J], 系统工程学报, 2009, 24(3):276-292,349.

[13]陆静. 基于分块极大值模型的商业银行操作风险计量研究[J]. 管理工程学报, 2012, 26(3): 136-145.

[14]陆静, 张佳. 基于极值理论和多元 Copula 函数的商业银行操作风险计量研究[J]. 中国管理科学, 2013,21(3): 11-19.

[15]明瑞星, 谢铨. 尾相关 Copula 在操作风险计量中的应用[J]. 统计与决策, 2013(1): 86-88.

[16]钱艺平, 林祥, 陈治亚. BMM 模型在商业银行操作风险度量中的应用[J], 统计与决策, 2010(10): 68-71.

[17]司马则茜,蔡晨,李建平. 度量银行操作风险的 POT 幂律模型及其应用[J]. 中国管理科学, 2009, 17(1),36-41.

[18]司马则茜,蔡晨,李建平. 基于 g-h 分布度量银行操作风险[J], 系统工程理论与实践, 2011(12): 2321-2327.

[19]宋坤, 宋鹏. 操作风险经济资本的度量方法与实证[J]. 统计与决策,2011, 340(16): 142-145.

[20]宋坤, 刘天伦. 小样本下贝叶斯参数估计法对操作风险的度量[J]. 统计与信息论坛, 2012,27(8): 27-32.

[21]王宗润, 汪武超, 陈晓红, 等. 基于 BS 抽样与分段定义损失强度操作风险度量[J]. 管理科学学报, 2012, 15(12): 58-69.

[22]吴恒煜,赵平,严武,等.运用 Student T – Copula 的极值理论度量我国商业银行的操作风险[J].运筹与管理,2011(1): 157 – 163.

[23]张宏毅,陆静.运用损失分布法的计量商业银行操作风险[J].系统工程学报,2008,23(4): 411 – 416.

[24]周艳菊,彭俊,王宗润.基于 Bayesian – Copula 方法的商业银行操作风险度量[J].中国管理科学,2011,19(4): 17 – 25.

[25]高丽君,郭艳丽.中国商业银行内部欺诈影响因素分析[J].山东财政学院学报,2011(4):23 – 27.

[26]赵蕾.保险企业操作风险度量研究[D].上海:同济大学博士学位论文,2007.03.

[27]DE F P, DEJESUS – RUEFF V, JORDAN J S, ROSENGREN E S. Capital and Risk: New Evidence on Implications of Large Operational losses[J]. Journal of Money, Credit & Banking, 2006, 38(7): 1819 – 1846.

[28]刘宗国.基于 ANP 方法的我国商业银行操作风险管理研究[D].青岛:青岛大学硕士学位论文,2009.06.

[29]EMBRECHTS P, FURRER H, KAUFMANN R. Quantifying Regulatory Capital for Operational Risk[R]. Working Paper, RiskDay, 2003. 10.

A Sample Size Study of External Data Measuring China Commercial Bank Operational Risks

GAO Lijun¹, SONG Hankun^{2,3}

- (1. School of Business Administration, Shandong University of Finance and Economics, Jinan 250014, China;
2. School of Business, Qingdao University, Qingdao 266071, China;
3. Shandong Road Sub – branch, Industrial and Commercial Bank of China Qingdao Branch, Qingdao 266000, China)

Abstract:With the Basel Capital Accord II and III successively issued, China Banking Regulatory Commission (CBRC) requires that China domestic banks have the provisions for operational risk capital. Operational risk capital provision needs accurate measurement of risk, and Chinese scholars have no choice but to use external data for the measurement because of the great difficulty in obtaining internal data. Now nobody in China can answer such questions as how much external data can reflect the overall characteristics and how much external data is needed at a minimum. By using non – parameter variable scale bootstrap method, selecting sample size through setting the mean quantile cut – point variation rate, and simulating annual operational risk capital, this paper concludes after contrasts and analyses that the simulation results are relatively stable when the minimum sample size is 800 cases (including extreme values). This measurement result is reasonable compared with the judgment of the 2013 operational risk capital in Chinese banking industries by Wind Data, which indicates that a measurement without considering whether the collected external data can reflect the overall characteristics may lead to some deviations in the result.

Keywords:commercial bank; operational risk; external data; sample size; variable scale bootstrap method

(责任编辑 刘小平)