

Logistic 模型与 Probit 模型 用于上市公司财务预警的比较

白承彪

(复旦大学经济学院 上海 200433)

【摘要】 本文运用 Logistic 模型与 Probit 模型对 2007~2009 年沪深两市制造业 57 家首亏公司及其配对公司进行对比分析, 结果发现二者都能对上市公司逐步恶化的财务状况进行良好预测, 但 Logistic 模型的预测效果要优于 Probit 模型。

【关键词】 财务困境 Logistic 模型 Probit 模型 ST 公司

财务困境又称“财务危机”或“财务失败”, 国内外对此尚无统一的定义。目前对此的共识是财务困境本质上就是财务危机, 最严重的财务困境就是企业破产。深入研究财务困境的成因, 探索其形成规律, 进而采取预警措施对企业来说是非常重要的。

国外学者通常将企业提出破产申请作为企业陷入财务危机的标志, 但目前阶段, 上市资格在我国仍然是一种珍贵的“壳”资源, 即使上市公司面临破产危险, 也会有其他企业将其接收(即借壳上市), 所以企业申请破产的可能性很小。然而, 将上市公司被特别处理视为陷入财务困境可解决这一概念界定问题。

亏损公司(LOSS)与营业外收入占利润总额比例(ZD1)之间不显著负相关, 假设 4 未能得到完全验证。实行新会计准则后, 亏损公司(LOSS×NAS)与 ZD1 之间虽然呈正相关关系, 但极端不显著。这说明实行新会计准则后, 亏损公司少计营业外收入以调减利润、进行盈余管理的假设不能得到验证, 即假设 4 在执行新会计准则后未能得到验证。

模型(2)的结果表明: 实行新会计准则后, 扭亏公司(NK×NAS)、小额盈利公司(WL×NAS)与剔除非流动资产处置净损益后的营业外收入占利润总额的比例(ZD2)之间呈显著的正相关关系。由于 ZD2 剔除了营业外收入传统的重要组成部分即非流动资产处置净损益, 从而进一步说明新债务重组准则被这些有强烈盈余管理动机的企业所利用。假设 3、假设 5 得到验证。

实行新会计准则后, 亏损公司(LOSS×NAS)与剔除非流动资产处置净损益后的营业外收入占利润总额的比例(ZD2)之间不显著负相关。假设 4 未能完全得到验证。

四、研究结论

第一, 债务重组收益的高低及其对利润总额的影响直接受会计准则规定的影响, 债务重组收益在准则规定可以计入利润的年份明显高于不能计入利润的年份。相对于前一个债务重组收益可以计入利润的阶段, 2007、2008 年度的债务重

一、研究方法

1. Logistic 模型。考察一个标准的线性回归方程: $\hat{y} = \alpha + \sum_{i=1}^n \beta_i x_i$, 如果对二分变量进行拟合, 则实际上拟合的是发生的概率, 参照上式可将二分变量回归方程表示为: $p = \alpha + \sum_{i=1}^n \beta_i x_i$ 。

但是, 进行线性回归所得估计值可能远离 [0, 1]。此外, 通常情况下因变量 p 与自变量 x_i 并非呈线性关系, 而是呈 S 型曲线关系, 这说明不能直接通过线性回归对二分变量进行拟合。然而, 对因变量进行 Logit 变换便可解决这两个问题。通过这种变换, Logit(p) 不但在 $(-\infty, +\infty)$ 取值, 而且 Logit(p) 往往和因

组收益对利润的贡献更大。

第二, 扣除偶然发生的营业外收入, 营业外收入最重要的两个组成部分就是债务重组收益和非流动资产处置净损益。在债务重组收益不能计入利润的期间, 上市公司大量利用了非流动资产处置净损益调节利润; 在债务重组收益不能计入利润的期间, 如 2007、2008 年度, 债务重组收益成为营业外收入的主要来源, 上市公司大量利用了债务重组收益调节利润。

为避免上市公司利用债务重组收益进行盈余管理, 笔者建议在遵循国际惯例的前提下, 将一次债务重组收益分期计入利润。这种改进可以在很大程度上减小债务重组对上市公司扭亏的贡献, 避免利润的大起大落, 客观反映上市公司的经营绩效。

【注】 本文系教育部人文社科基金资助项目“基于会计准则变迁的上市公司盈余管理研究”(项目编号: 09XJA790006) 的阶段性研究成果。

主要参考文献

1. 谢德仁, 张高菊. 金融生态环境、负债的治理效应与债务重组——经验证据. 会计研究, 2007; 12
2. Burgstahler D., Dichev I.. Earnings Management to Avoid Earnings Decrease and Losses. Journal of Accounting and Economics, 1997; 24

变量呈线性相关关系。将 $\text{Logit}(p)$ 作为因变量对自变量 x_i 进

行回归,便得到 Logistic 回归模型: $\text{Logit}(p) = \alpha + \sum_{i=1}^n \beta_i x_i$ 。将模

$$p = \frac{\exp\left(\alpha + \sum_{i=1}^n \beta_i x_i\right)}{1 + \exp\left(\alpha + \sum_{i=1}^n \beta_i x_i\right)}$$

型逆推可得: $p = \frac{\exp\left(\alpha + \sum_{i=1}^n \beta_i x_i\right)}{1 + \exp\left(\alpha + \sum_{i=1}^n \beta_i x_i\right)}$ 。
 2. Probit 模型。Probit 模型与 Logistic 模型一样,都是定性回归模型的一种,都可用于处理二变量回归问题。Probit 模型产生的原因与 Logistic 模型的相同,只不过其在解决线性回归所带来的两个难题上不同于 Logistic 模型。对于同样的线性回归方程: $\hat{y} = \alpha + \sum_{i=1}^n \beta_i x_i$, 为确保 y 发生的概率介于 0 和 1 之

间,且使事件发生的概率模型为非递减函数,该模型采取概率分布函数将 y 做转换。Probit 分布函数的形式为: $p = F(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ 。与 Logistic 模型相同的是,Probit 模型将临

界概率值(通常为 0.5)作为事件发生与否的判断标准,若事件发生的概率大于临界值,则判定事件发生,反之亦然。

二、研究设计

1. 样本选取。首先,为避免出现年度效应和高估模型的预测能力,选择近三年来新增的被 ST 公司(排除因非经营性原因、偶发事件等被 ST 者),并按照同时期、同行业、同规模(资产比不超过 15%)的原则选择未被 ST 者作为配对样本,将全部样本分为训练样本和预测样本。其次,从时序角度考察公司是否陷入财务困境。假定上市公司被 ST 年份为第 T 年,分别用第 T-1、T-2、T-3 年的数据所得到的模型来判断所有公司第 T 年是否被 ST。据此,我们从中国经济金融数据研究库选取了沪深两市 2007~2009 年制造业中的 57 家 ST 上市公司和 57 家配对公司共 114 家公司作为总样本,用其中的 94 家作为训练样本估计模型,用 20 家作为预测本来检验模型的稳健性。

2. 指标选取。结合国内外研究成果,本文初步选取了 27 个指标(见表 1),这些指标全面反映了企业的盈利能力、发展能力、偿债能力和营运能力。

虽然根据表 1 中的指标可以对上市公司财务状况做出较为全面的评价,但是由于选取的指标比较多,增强了分析的复杂性,并且这些指标反映的信息存在一定的重叠。因此,必须选出那些最能区分 ST 公司和非 ST 公司状况的指标。下面的数据处理皆运用 SPSS16.0 软件完成。

表 1 初步选取的财务指标

每股收益	每股净资产	每股营业收入	净资产收益率	总资产增长率	总资产增长率	营业收入增长率	营业利润增长率	流动比率	速动比率	存货流动负债比率	现金流动负债比率	资本充足率	现金负债比率	债务资本比率	债务资产比率	存货周转率	应收账款周转率	流动资产周转率	资产周转率	存货销售期	应收账款回收期	市盈率	市净率	市销率	托宾 O	
X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	X26	X27

3. 指标筛选。

(1)正态性检验。由于训练样本的数量大于 50,因此采用 Kolmogorow-Smirnov Z 统计量在 95%的置信水平上对预选指标进行正态性检验。检验结果表明,第 T-1 年有 X2、X17、X20 不能拒绝服从正态分布;第 T-2 年有 X1、X14、X17、X20、X27 不能拒绝服从正态分布;第 T-3 年有 X2、X7、X14、X17、X18、X20、X27 不能拒绝服从正态分布。因此,三年的数据总体上来说并不服从正态分布。

(2)均值差异检验。由于三年的数据总体上不服从正态分布,因此只能对这些初始指标的均值进行非参数检验,以考察哪些指标能够显著区分 ST 公司和非 ST 公司。在 95%的置信水平上,用 Kruskal-Wallis 统计量进行检验,各年差异显著变量汇总如表 2 所示:

年份	各年差异显著变量
T-1	X1、X2、X4、X5、X6、X7、X8、X9、X10、X11、X12、X13、X14、X15、X16、X17、X24、X27
T-2	X1、X4、X5、X6、X9、X10、X11、X13、X14、X15、X16、X17、X24、X27
T-3	X1、X4、X11、X14、X16、X17

从表 2 可以看出,X1、X、X11、X14、X16 和 X17 在三年中的差异都显著。因此,从各年变量选取的统一角度来讲,可以选取这 6 个变量作为进一步研究的对象。

(3)自相关检验。因为研究的变量整体不服从正态分布,所以可用 Spearman 相关系数检验变量间的相关性。检验结果表明,变量 X1、X4 之间,变量 X14、X16、X17 之间在三年内皆高度相关。保留与其他变量相关系数小的变量,剔除与其他变量相关程度高的变量,对高度相关的变量进行逐个剔除,同时考虑到能最大程度地提高模型的总体预测能力,最终保留的变量为 X1、X11、X16。

(4)多重共线性检验。多重共线性检验主要使用容忍度(TOL)和方差膨胀系数(VIF)两个指标作为检验标准。容忍度小于 0.1 即方差膨胀系数大于 10,则说明多重共线性很强。各变量三年的容忍度和方差膨胀系数如表 3 所示:

	T-1		T-2		T-3	
	TOL	VIF	TOL	VIF	TOL	VIF
X1	0.807	1.239	0.759	1.317	0.956	1.047
X11	0.829	1.206	0.818	1.222	0.941	1.063
X16	0.967	1.034	0.745	1.343	0.913	1.095

从表 3 可以看出,X1、X11、X16 并不存在多重共线性,且反映了上市公司的偿债能力、盈利能力这两个最重要的方面,能够对三年的 ST、非 ST 公司进行显著区分。

三、实证分析

1. Logistic 模型。进行后向逐步回归,结果如下:

表 4 Logistic 回归模型系数汇总

年份		B	S.E.	Wald	Df	Sig.	Exp(B)
T-1	X1	-12.251	2.842	18.585	1	0.000	0.000
	Constant	-1.388	0.475	8.536	1	0.003	0.249
T-2	X1	-3.687	1.068	11.919	1	0.001	0.025
	X16	0.638	0.309	4.262	1	0.039	1.893
T-3	Constant	-0.950	0.477	3.960	1	0.047	0.387
	X1	-3.034	1.382	4.820	1	0.028	0.048
T-3	X16	1.170	0.378	9.578	1	0.002	3.221
	Constant	-1.081	0.502	4.636	1	0.031	0.339

各年的回归函数为:

$$\text{第 T-1 年: } p = \frac{\exp(-1.388 - 12.251X1)}{1 + \exp(-1.388 - 12.251X1)}$$

$$\text{第 T-2 年: } p = \frac{\exp(-0.950 - 3.687X1 + 0.638X16)}{1 + \exp(-0.950 - 3.687X1 + 0.638X16)}$$

$$\text{第 T-3 年: } p = \frac{\exp(-1.081 - 3.034X1 + 1.170X16)}{1 + \exp(-1.081 - 3.034X1 + 1.170X16)}$$

Logistic 模型的预测结果见表 5:

表 5 Logistic 回归模型预测结果

年份	原始值	训练样本	准确率	预测样本	准确率
T-1		0	1	0	1
	0	43	4	91.5%	9
	1	5	42	89.4%	2
总体准确率			90.4%		85.0%
T-2		0	1	0	1
	0	43	4	91.5%	7
	1	13	34	72.3%	2
总体准确率			81.9%		75.0%
T-3		0	1	0	1
	0	33	14	70.2%	7
	1	13	34	72.3%	3
总体准确率			71.3%		70.0%

2. Probit 模型。Probit 模型的回归结果见表 6:

表 6 Probit 回归模型系数汇总

年份	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
T-1	X1	-6.528	1.288	-5.069	0.000	-9.053	-4.004
	Intercept	-0.809	0.260	-3.110	0.002	-1.069	-0.549
T-2	X1	-1.799	0.465	-3.873	0.000	-2.710	-0.889
	X16	0.417	0.177	2.355	0.019	0.070	0.764
T-3	Intercept	-0.609	0.281	-2.165	0.030	-0.890	-0.328
	X1	-1.834	0.800	-2.292	0.022	-3.403	-0.266
T-3	X16	0.724	0.225	3.216	0.001	0.283	1.165
	Intercept	-0.679	0.301	-2.257	0.024	-0.980	-0.378

表 6 中,各年的回归函数为:

$$\text{第 T-1 年: } p = F(-0.809 - 6.528X1)$$

$$\text{第 T-2 年: } p = F(-0.609 - 1.799X1 + 0.417X16)$$

$$\text{第 T-3 年: } p = F(-0.679 - 1.834X1 + 0.724X16)$$

Probit 模型的预测结果见表 7:

表 7 Probit 回归模型预测结果

年份	原始值	训练样本		准确率	预测样本		准确率
		0	1		0	1	
T-1		0	1		0	1	
	0	43	4	91.5%	9	1	90.0%
	1	5	42	89.4%	2	8	80.0%
总体准确率				90.4%			85.0%
T-2		0	1		0	1	
	0	43	4	91.5%	6	4	60.0%
	1	12	35	74.5%	2	8	80.0%
总体准确率				83.0%			70.0%
T-3		0	1		0	1	
	0	33	14	70.2%	6	4	60.0%
	1	15	32	68.1%	3	7	70.0%
总体准确率				69.2%			65.0%

四、结论

比较 Logistic 模型与 Probit 模型的预测结果可以看出:对于训练样本,Logistic 模型在 T-1、T-2、T-3 年的总体预测准确率分别为 90.4%、81.9%、71.3%;Probit 模型在 T-1、T-2、T-3 年的总体预测准确率分别为 90.4%、83%、69.2%。

这说明,一方面随着 ST 年份的临近,两个模型的预测准确率也在提高,这是符合上市公司财务状况逐步恶化事实的;另一方面,对 T-1 年而言,使用任何一种模型对训练样本进行预测都是无差别的。此外,Logistic 模型对训练样本在第 T-2 年的预测准确率要低于 Probit 模型对训练样本在第 T-2 年的预测准确率。

事实上,这种差别和第 T-1 年的差别并不重要。这是因为大多数上市公司被 ST 就是由于这些上市公司前两年连续亏损,在这种情况下,我们用前两年的数据来预测这些上市公司第 T 年是否被 ST 就没有太大的意义,而有效预测往往以第 T-3 年为起点。很明显,在第 T-3 年,Logistic 模型的预测准确率要高于 Probit 模型的预测准确率。这也是实践中大多数研究者采用 Logistic 模型预测企业信用风险的一个重要原因。

主要参考文献

1. 陈静. 上市公司财务恶化预测的实证研究. 会计研究, 1999; 4
2. 吴世农, 卢贤义. 我国上市公司财务困境的预测模型研究. 经济研究, 2001; 6
3. 林震岩. 多变量分析——SPSS 的操作与应用. 北京: 北京大学出版社, 2007
4. 张文彤, 阎洁. SPSS 统计分析高级教程. 北京: 高等教育出版社, 2004