



清华大学全球证券市场研究院  
Institute for Global Securities Market, Tsinghua University



Asian Institute of  
Digital Finance

# 公司违约预测模型 详细介绍与数据来源



清华大学全球证券市场研究院

新加坡国立大学亚洲数字金融研究所

## 一、 算法与模型介绍

### CRI 公司违约预测模型<sup>1</sup>

违约概率（PD）是新加坡国立大学信用研究计划（NUS-CRI）的核心信用指标，由该机构研发的公司违约预测模型计算，该模型衍生于由段锦泉教授团队所构建的远期密度模型<sup>2</sup>（Forward Intensity Model）。该模型能够对宏观金融和企业微观数据进行动态学习，从而生成具有前瞻性的上市公司违约概率期限结构。

CRI 公司违约预测模型的基础是远期条件概率（Conditional Forward Probability），如下式所示：

$$p_{i,t}(\tau) = P_{\tau}(X_t, Y_{i,t})$$

上式中 $p_{i,t}(\tau)$ 表示公司 $i$ 当前处于 $t$ 时刻，在 $t$ 至 $t + \tau$ 时期存活，并在 $t + \tau$ 至 $t + \tau + 1$ 时期违约的概率。 $P_{\tau}(X_t, Y_{i,t})$ 表示远期密度函数（Forward Intensity Function），其中 $X_t$ 表示宏观金融风险因子（Macro-financial Risk Factors）， $Y_{i,t}$ 表示公司层面特征（Firm-specific Attributes）。上述协变量 $X_t$ 和 $Y_{i,t}$ 的具体信息如下表：

表 1 CRI 公司违约预测模型输入协变量

	模型输入变量	描述
宏观金融风险因子	股票指数回报率	主要股票市场指数年回报率，经缩尾和货币调整
	短期无风险利率	3 月期国债收益率
	金融公司违约距离 (Distance-To-Default)	金融/非金融企业的违约距离中位数
	非金融公司违约距离	
公司层面特征	违约距离 (水平)	波动率调整杠杆 (Volatility-adjusted Leverage)，CRI 在 Merton (1974) 基础上进行了特殊处理
	违约距离 (趋势)	
	现金/总资产 (水平)	针对金融公司
	现金/总资产 (趋势)	
	流动资产/流动负债 (水平)	针对非金融公司
	流动资产/流动负债 (趋势)	
	净收入/总资产 (水平)	盈利能力
	净收入/总资产 (趋势)	
	相对规模 (水平)	每一家公司的市值与过去一年经济体市值中位数的比率的对数
	相对规模 (趋势)	
	相对市值账面比 (Market-to-Book Ratio)	单个公司的市场估值错误/未来增长机会相对于经济体市净率中位数的比值
	异质性波动率 (Idiosyncratic Volatility)	市场模型残差的标准差
	国企虚拟变量	根据中国政府公开的企业官方信息获取一家公司的国有企业属性信息，并定期更新国有企业名单。

此外，中国 A 股上市公司具有一定的特殊性。市场普遍认为，中国的国企较非国企更“安全”。该模型考虑到了中国国企与非国企之间的差异，在中国模型中加入了国企虚拟变量<sup>3</sup>。图 A-1 展示了旧模型与加入国企虚拟变量的新模型对于国企和非国企违约数量的预测结果，可以看出，新模型更加适用于我国的公司违约风险评估。

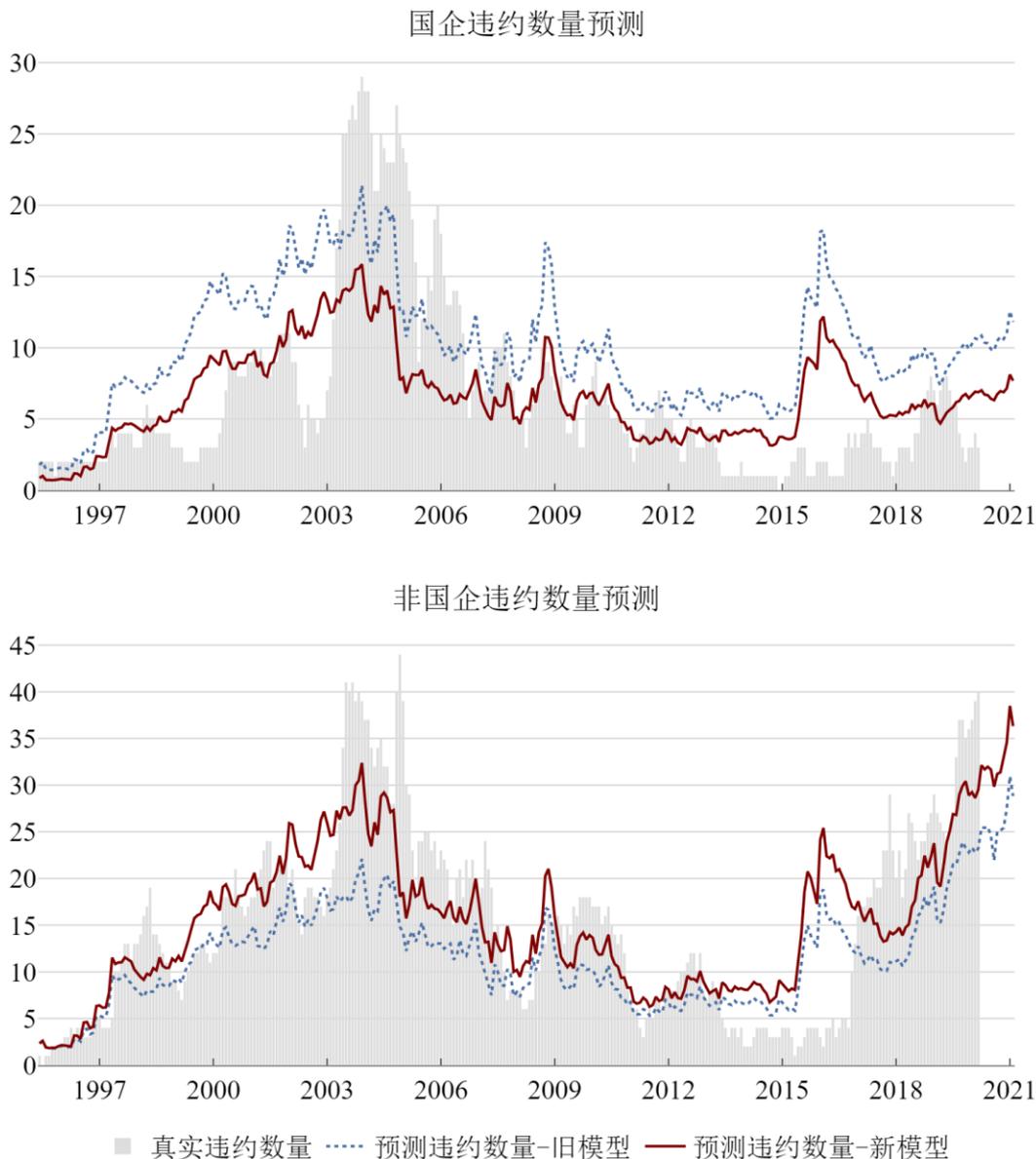


图 1 国企与非国企违约数量预测

基于  $p_{i,t}(\tau)$ ，通过典型的生存-退出方程（Survival-exit Formula），可以得到不同期限结构的多期违约概率。CRI 对潜在的远期密度函数进行参数化，并根据 CRI 数据库中更新的新信息，逐月估计每个校准组的参数。

在模型表现的检验方面，准确率（Accuracy Ratio）是一种常用

的定量度量方法，用于评估违约预测模型的预测能力。该指标是 (a) 检验模型表现与随机模型表现之差与 (b) 完美模型与随机模型表现之差的比值。对于一个完全无信息的模型，其准确度指标为零。准确率是指，如果违约公司在违约前被分配到信用风险最高的组别中，那么模型就正确地区分了安全公司和风险公司。CRI 公司违约预测模型在其覆盖的所有地区和经济体中都获得了较高的准确率，表明其表现良好。

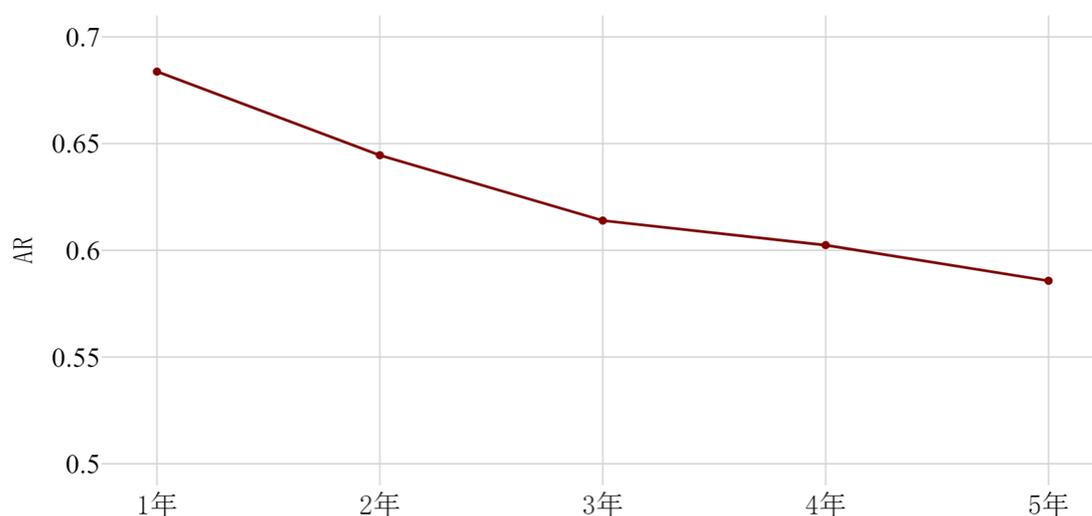


图2 我国 A 股上市公司违约预测能力

图2展示了CRI公司违约预测模型对中国公司违约的预测能力，模型对于未来1年以及5年内公司违约的预测能力分别高于0.6和0.5。虽然对于较长期限的违约事件本模型预测能力有所下降，但仍然具有一定的预测力。本文的分析主要基于1年累计违约概率指标，其预测准确度为0.685，即保证了模型的预测能力，也为投资者和监管者提供了足够长的预测期限。

为了充分展示本模型的预测能力，我们以 2020 年 11 月天齐锂业的违约事件为例，分析了 PD 对于中国典型违约事件的预测能力<sup>4</sup>。天齐锂业股份有限公司（简称天齐锂业）是中国最大的锂生产企业，也是世界上最大的锂生产商之一，控制着全球超过 46% 的锂产量。从历史来看，天齐锂业 PD 自 2020 年 3 月一路攀升，于 2020 年 10 月 PD 达到峰值 471.58，最终于 2020 年 11 月违约。从远期 PD 来看，自违约事件发生前 1 年起，远期 PD 就预测其违约风险将在未来四年内走高。因此，PD 对于中国典型违约事件具有很强的预测作用。

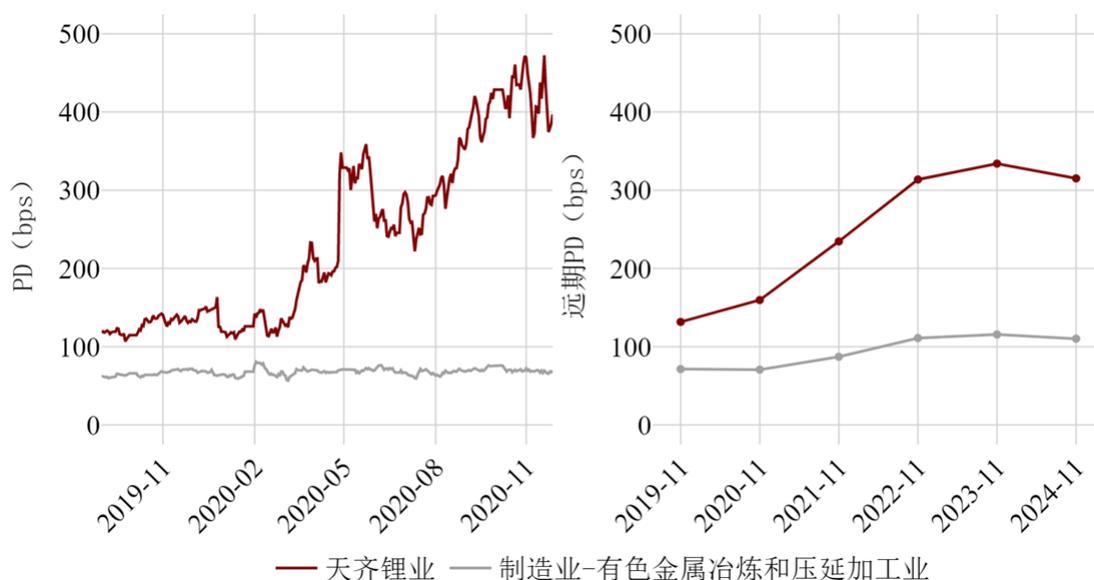


图 3 中国市场违约事件预测能力展示——以天齐锂业为例

<sup>1</sup> 算法与模型均基于 NUS-CRI 违约概率白皮书

<https://d.nuscricri.org/static/pdf/Probability%20of%20Default%20White%20Paper.pdf>

<sup>2</sup> 详见 Duan, J.-C., J. Sun, and T. Wang, 2012, Multiperiod Corporate Default Prediction - A Forward Intensity Approach. Journal of Econometrics 170, 191-209.

<sup>3</sup> 具体参见 NUS-CRI 技术报告 <http://d.nuscricri.org/static/pdf/2020update2Addendum2.pdf>

<sup>4</sup> NUS-CRI 2020 年 2 月 4 日-10 日的每周信用简报详细分析了天齐锂业信用违约风险快速上升的原因 <https://nuscricri.org/media/static/images/thumbnail-pdf/WCBFEB04FEB102020.pdf>

## 违约概率隐含评级<sup>5</sup>

违约概率隐含评级（PDiR）由 CRI 于 2011 年提出，提供了一个方便、直观的信用评级。具体来说，该模型将公司信用质量通过违约概率指标映射到了信用评级。PDiR2.0 开发于 2020 年，并发表与 Duan 和 Li（2021）<sup>6</sup>。该方法基于标准普尔或穆迪全球企业评级池的历史信用评级迁移信息，而不是仅仅依赖报告的违约率。

**表 2 违约率映射**

	标普公布的违约率	PDiR2.0 已实现的违约率	国内评级已实现的违约率
AAA	0.0000%	0.0552%	0.4663%
AA	0.0147%	0.0282%	0.7173%
A	0.0516%	0.0401%	0.8334%
BBB	0.1676%	0.1324%	1.5762%
BB	0.7479%	0.7965%	2.8216%
B	3.7477%	4.6090%	4.2125%
CCC	27.6436%	26.8206%	4.6690%
CC	50.5455%	56.5957%	2.4342%
C	50.0000%	66.6667%	5.7504%

资料来源：NUS-CRI

表 2 汇报了各信用评级对应的标普公布的违约率、PDiR2.0 已实现的违约率。此外，我们还根据国内评级机构<sup>7</sup>的信用评级信息计算了每个评级所对应的违约率。可以看出，PDiR2.0 违约率与标普较为接近，而国内评级违约率与标普相差较大。表 3 汇报了标普信用评级对于各评级的详细说明。根据标普的定义，BBB 及以上评级定义为投资级，BB 及以下定义为投机级。

表 3 标准普尔信用评级介绍<sup>8</sup>

评级符号	评级说明
AAA	偿还债务能力极强（Extremely strong）
AA	偿还债务能力很强（Very strong）
A	偿还债务能力强（Strong），但略微易受外在环境及经济状况变动的不利因素所影响
BBB	具有适当（Adequate）偿债能力，但偿债能力较可能因不利经济状况而减弱
BB	短期内不太容易受到影响，但持续存在的重大不稳定因素，或不利的商业、金融、经济状况，可能导致发债人没有足够能力偿还债务。
B	较易受到不利的商业、金融和经济条件的影响，但目前有能力偿还债务
CCC	目前有可能违约，发债人能否履行财务承诺将取决于商业、金融、经济条件是否有利。
CC	违约的可能性高。违约尚未发生，但预计会发生。
C	目前违约的可能性高，且最终违约追偿比率预计会低于其他更高级别的债务。
D	发债人未能按期偿还债务，或违反推定承诺；也用于已提交破产申请的情况。

资料来源：标准普尔信用评级

<sup>5</sup> 算法与模型均基于 NUS-CRI 违约预测隐含评级模型白皮书

[https://d.nuscri.org/static/pdf/PDiR2.0\\_White\\_Paper\\_2021.pdf](https://d.nuscri.org/static/pdf/PDiR2.0_White_Paper_2021.pdf)。

<sup>6</sup> Duan, J. C., & Li, S. (2021). Enhanced PD-implied ratings by targeting the credit rating migration matrix. *The Journal of Finance and Data Science*, 7, 115-125.

<sup>7</sup> 包括中诚信、联合和大公信用评级机构。

<sup>8</sup> 标准普尔的信用评级标准在中国的适用性有限，仅供参考。

## 二、名词解释

名称	定义	来源
违约/违约类型	违约事件可分为以下三类：1、破产申请、接管、管理、清算或因及时清算利息和/或本金支付而陷入的任何其他法律困境；2、错过或延迟支付利息及/或本金，但在宽限期内延迟支付的除外；3、债务重组/不良债务交换，即向债务持有人提供新的证券或一揽子证券以偿还债务（例如，将债务转换为股权、较低息票或面值的债务、较低优先级的债务、较长期限的债务）。	NUS-CRI
PD	即预测累计违约概率（Probability of Default，简称PD），预测期限默认为1年（即未来一年内发生违约的概率），单位默认为基点（bps），报告中多用（%）表示（1%=100bps）。模型介绍与预测能力检验结果于附录 A.1 中汇报。	NUS-CRI
远期 PD	即远期违约概率，由 CRI 计算，用于估计公司在未来一段时间内的信用风险。例如，第二年远期违约概率即未来1年末至2年末期间的预测累计违约概率，其中假设公司在未来1年内存活。	NUS-CRI
远期密度模型	即 Forward Intensity Model，由段锦泉教授等人建立，详见附录 A.1	NUS-CRI
PD <sub>iR</sub>	将违约概率（PD）与标普评级相对应，转换为更加便于监测的字母评级。详见附录 A.1。	NUS-CRI
（信用）违约风险	是指借款人、证券发行人或交易对方因各种原因，不愿或无力履行合同条件而构成违约，致使银行、投资者或交易对方遭受损失的可能性。	
真实违约数量	真实发生违约的公司数量，默认为当季（日）未来一年内真实发生违约的公司数量，若标注“当季值”则表示当季内真实发生违约的公司数量。	
真实违约（概）率	真实发生违约的公司数量占同期全部公司数量的比值，单位为百分比（%），默认为当季（日）未来一年内真实发生违约的公司占比，若标注“当季值”则表示当季内真实发生违约的公司占比。	

### 三、数据与资料来源

数据与资料	数据与资料来源
PD	NUS-CRI
PDiR	NUS-CRI
违约数据	多种国际数据分销商（如汤森路透 Datastream、彭博数据库等）以及上市公司公告等。
案例相关资料	上市公司财报、公告，以及财经类新闻媒体（新浪财经、搜狐财经、经济观察网、每日经济新闻等）。
上市公司数据	万得数据库、彭博数据库、国泰安数据库、CCER 中国经济金融数据库等
实际 GDP 增速	国家统计局
消费者价格指数（CPI）	国家统计局
生产者价格指数（PPI）	国家统计局
上证综指	上海证券交易所
国债收益率	中国债券信息网
银行间同业拆借利率	中国人民银行
货币供应量（M2）同比增速	中国人民银行
VIX 指数	芝加哥期权交易所
人民币实际汇率指数	国际清算银行



## 清华大学全球证券市场研究院



地址：北京市海淀区清华大学经济管理学院伟伦楼

邮编：100084

邮箱：[igsm@sem.tsinghua.edu.cn](mailto:igsm@sem.tsinghua.edu.cn)

网址：<http://igsm.tsinghua.edu.cn>



清华大学全球证券市场研究院

Institute for Global Securities Market, Tsinghua University