

国内外商业银行压力测试研究综述

彭志慧

(西南财经大学中国金融研究中心, 成都, 610074)

摘要: 压力测试是针对重大风险而发展起来的一门风险管理技术, 旨在测度银行在极端情景下的抗压能力。本文首先从国外压力测试研究现状出发, 分析了早期压力测试框架, 以及压力测试技术的最新进展; 其次, 阐述了国内宏观压力测试和信用风险压力测试两类研究; 最后评述国内外的研究现状。

关键词: 压力测试 商业银行 风险管理 信用风险

压力测试的发展可以大致被分为两个时期: 2007 年以前的压力测试和 2007 年以后的压力测试。这样的划分是因为, 2007 金融危机的爆发揭露了风险管理包括压力测试技术上的诸多不足, 实务界和学术界都开始了对压力测试的反思和完善工作, 因此 2007 年前后关于压力测试的研究差别比较大。当然, 这也是为了研究方便所做的一个大致划分, 并不绝对。

早期的研究以 FSAP 框架为基础, 其研究角度主要集中在压力测试框架探讨, 学者从识别脆弱性、构建情景、测度资产负债表变化、考虑次轮效应和结论解释等方面对压力测试框架进行了探讨。2007 金融危机之后, 出于对金融危机的反思, 这一时期的压力测试研究对情景的概率、情景生成模型、压力传导模型、压力测试实施以及影响测度等方面进行了深入研究。

一、国外压力测试研究现状

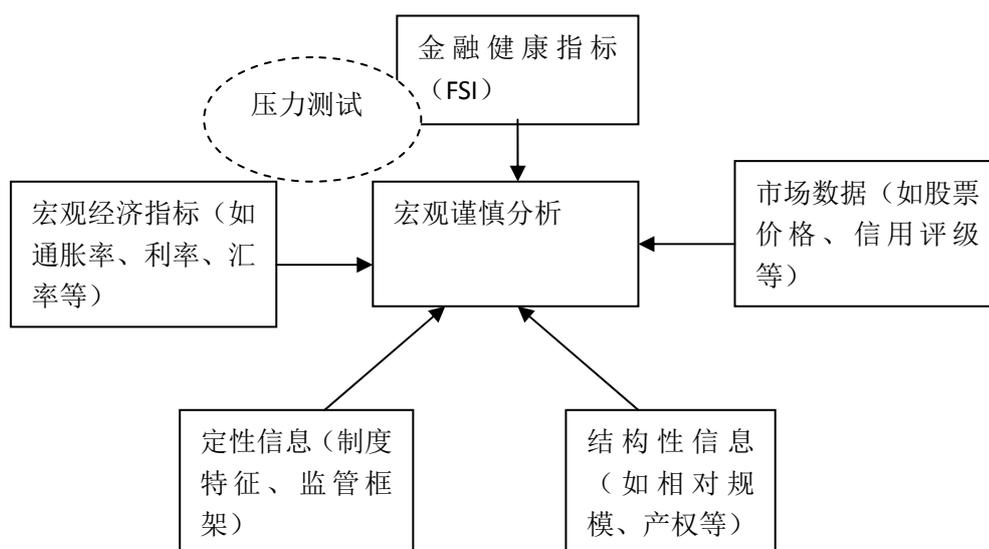
(一) 早期的压力测试: FSAP 框架

早起压力测试的发展同监管当局的推动有很大关系。早期各国监管当局实施的压力测试主要是在世界银行和国际货币基金组织发起的“金融部门评估规划”(Financial Sector Assessment Program, FSAP) 框架下进行的, 笔者将这一类压力测试方法称作“FSAP 框架”。

1999 年, 世界银行和 IMF 发起金融部门评估规划, 其目的评估成员国金融体系的强度和脆弱性, 亦即宏观谨慎分析 (macroprudential analysis)。¹宏观谨慎分析包括金融市场状况监测和宏观金融状况分析。它是宏观经济脆弱性评估框架的一个组成部分, 而宏观经济脆弱性框架包括资产负债表方法、债务可持续性分析 (debt sustainability analysis) 和宏观经济状况监测。这个框架的一个重要组成部分就是用压力测试方法进行定量分析。图 1.2 描述了宏观谨慎分析的基本框架。

图 1.2 宏观谨慎分析框架

¹ Evans et al.(2000), Sundararajan et al.(2002), International Monetary Fund(2003 a, b), and IMF-World Bank (2003)



资料来源：Sundararajan et al. Financial Soundness Indicators: Analytical Aspects and Country Practices. Occasional Paper No.212(Washington,DC:International Monetary Fund), 2002

Winfried Blaschke et al. (2001)、Matthew T. Jones et al. (2004) 详细论述了 IMF 压力测试的框架，以及 FSAP 国家在进行压力测试方面的经验教训。这一框架的主要程序包括：识别脆弱性；构建情景；测度资产负债表变化；分析次轮效应；结论解释。

(1) 识别脆弱性

因为对金融体系的每一风险因子进行压力测试是不太可能的，所以简化的方法是寻找那些可能给金融体系带来重大影响的关键脆弱性。可以使用宏观经济指标和金融健康指标 (Financial Soundness Indicators, FSIs) 来识别和理解金融体系在冲击面前的脆弱性以及金融体系吸收潜在损失的能力。表 1.2 是主要的金融健康指标。

表 1.2 核心金融健康性指标

资本充足性指标	监管资本/风险加权资产 核心资本/ 风险加权资产
资产质量指标	不良贷款/总贷款 不良贷款准备净值/资本 贷款部门分布/总贷款 大的风险敞口/资本
收益和盈利能力	资产收益率 股本收益率 非利息费用/总收入
流动性	流动性资产/总资产 流动性资产/短期负债
市场风险敏感性指标	资产久期 负债久期 外汇净公开头寸/资本

资料来源：Matthew T. Jones et al. Stress Testing Financial Systems: What to Do When the Governor Calls. IMF Working Paper, 2004

(2) 构建情景

一旦关键脆弱性已经确认,接下来要做的就是构建情景。这是压力测试中最基础的部分。这一阶段的内容包括使用数据和模型进行测试,分析金融体系在特定情境下对假定冲击的反应。

在理想情况下,宏观计量经济模型或模拟模型应构成压力测试情景的基础部分。这是因为系统性压力测试的一个目标是理解经济环境的重大变化对金融体系的冲击。使用宏观模型对分析金融体系和实体经济之间的联系提供了一个框架。只是很多国家往往缺乏合适的宏观经济模型。

Matthew T. Jones et al. (2004) 提供了一些参考性的情景。比如,当一经济体就业率上升、居民可支配收入上升、房价急剧上升以及低利率情况引发抵押贷款繁荣时,我们可以假设一个失业率上升、可支配收入下降以及利率急剧上升的情景。再如,假设宏观经济模型表明,由于大量资本流入导致汇率高估,相伴随的是信贷增长导致的建筑行业投资井喷。再对机构报表分析发现急剧增长的外币风险暴露,微观水平的 FSI 和个体银行报表信息表明贷款违约风险逐步上升。这时,一个可能的情景是:资本流动突然反转,汇率急剧贬值。

(3) 测度资产负债表变化

情景一旦设立,接下来要做的就是研究该情景下金融机构资产负债表的变化。有两种方法将情景转换或映射到资产负债表之中。一种是自下而上法(bottom-up approach),另一种是自上而下法(top-down approach)。

自下而上法使用个体资产组合的数据来估计冲击的影响,自上而下法则是使用总量数据来估计冲击的影响。采用自下而上法时,在一个情景下对各种冲击反应的估计是在资产组合水平下进行的,使用的是个体金融机构分散的数据。这种方法得出的结论是可以加总分析或比较分析的。这种方法的优点是可以较好的使用个体机构资产组合方面的数据。但个体机构在植入情景上的会有差异,这可能会产生一些不一致的情况。自上而下法通常用来估计一组金融机构对一个特定情景的反应。这种方法一般更容易实施,因为它仅需要总量数据,而且采用的是一致的方法。

采用自下而上方法进行系统性压力测试需要两次映射:第一次映射测度宏观情景对风险因子的影响;第二次映射测度风险因子对资产组合价值的影响。其过程是:情景——风险因子——不同金融工具价格——资产组合价值重估——资产负债表变化。

(4) 次轮效应(second-round effects)

大多数压力测试方法都假定,风险因子改变时,资产组合的行为没有改变,或资产组合结构不会被调整。如果时间期限相对较短,或资产组合的调整需要时间,那这一假设是有意义的。比如,假设大的贷款资产组合在1—3个月的时间范围内只能进行有限的调整,就可能是一个有意义的假设。这一假设对那些对金融体系或宏观经济影响小的机构而言也适用,因为反馈效应相对较小。

一旦情景或冲击的时间界限超过一年,假设没有反馈效应就不再合理。同样,对系统性重要的机构或整个金融体系而言,无反馈效应的假设就显得太简化了。在较长的期限内,政策环境可能会发生改变,监管当局可能会对冲击作出反应,金融安全网的缓冲效应也会发挥作用。

Hoggarth, Whitley (2003) 认为,次轮效应是很明显的:因为交易对手破产带来的信贷损失,因为受损银行增加的融资成本,因为资产组合调整对总需求的影响,等等。这些效应在压力测试期间可能存在,如何将其纳入压力测试是一个挑战。一种方法是使用传染模型(contagion model),这些模型视图估计关键机构失败对其他机构以致整个金融体系的影响。

(5) 结论解释

压力测试结论有助于评估关键变量发生重大变动的影 响，但是它提供的并非损失的精确测度。通常情况下，我们只知道损失的大小，而不知道发生这种损失的可能性有多大。这是压力测试的一个弊端。在解释压力测试结论时还要注意以下几点：（1）模型错误问题。如果压力测试所使用的模型并不正确，那压力测试的结论也是无效的。因此可以将压力测试结论同其他风险测度方法的结论进行比较，做交叉检查。（2）压力测试不太可能发现所有风险以（如操作风险和法律风险）及风险之间的相互作用，它给出的仅仅是机构风险的部分特征。（3）压力测试考虑的仅仅是银行收入来源的一部分，而银行可能还有其他收入没有被压力测试分析到。

（二）压力测试技术的最新进展

2007 金融危机爆发后，很多国家监管当局重新强调加强压力测试的必要性，并开始加强压力测试操作指引。监管当局采用的压力测试工具跟金融机构的有所不同，因为中央银行或监管当局更关注整个金融体系的稳定性。同时危机也暴露了金融机构风险管理技术包括压力测试技术的缺陷，引起了一轮压力测试技术改进浪潮。这里我们按照压力测试的操作流程对压力测试技术的最新进展做一个回顾。

（1）情景构建

压力测试要求构建极端但可能发生的情景。具体到操作层面有两个问题：一是情景的可能性（plausibility）的测度。压力测试的一个麻烦是，它研究的是小概率事件，如果这个事件发生的可能性太低，那压力测试的意义就不大。二是情景的严峻度（severity）。如果情景的发生不会对机构造成重大影响，那压力测试也没有多大意义。因此，如何寻求极端但可能事件就是一个问题。

早期的一些学者已经意识到对压力测试情景设定概率的重要性。如果我们不知道情景发生的可能性，那我们就无法判断压力测试结论的重要性（Berkowitz, 1999）。问题是，压力情景通常都是历史上很少发生的事件，或者从来就没有发生过，这样要为其设定概率就很困难。通常的作法是采取主观判断法，即选用主观概率，如 FSAP 框架下的专家判断法。但这种方法主观性太强。不同风险管理者或在不同时期，即使是采用同一情景也可能会出现不同的结果。

一种方法是利用 VAR/VECM 框架生成具有概率分析的压力情景。冲击的规模由自回归序列中无条件标准离差确定，在正态分布假设下，冲击的规模就会对应一个概率。这类模拟的尾部就代表极端情景。这类情景同宏观结构模型的区别在于，它仅仅基于一个概率方法，没有经济学理论支撑。FSAP 框架则相反，它使用的是宏观计量经济结构模型。这种方法下，模型生成的点估计值有一个单一的未来路径，这个路径是压力情景下的条件平均路径，没有概率分析。

Pesaran、Schuermann、Treutler 和 Weiner（2006）最早使用 VAR 模型生成一个概率情景来分析信用风险。作者使用冲击响应函数来检测对一个宏观经济变量单独的冲击如何影响其他变量。冲击反应函数假定其他变量根据它们同被冲击变量的历史协方差被替换，因此这种方法恰当的考虑了冲击间的相互关系。这种方法对特定问题风险管理特别有价值，特别的，它可以产生一个可能的冲击情景等级序列。

Jiménez、Mencía（2007, BoS）和 Castren、Fitzpatrick、Sydow（2008, ECB）以及日本银行在压力测试中采用了这种概率方法生成情景。在 BoJ 的模型中，压力测试估计一个负的 GDP 冲击的影响，冲击的大小对应的概率为 1%。Jiménez 和 Mencía 采用一个 GDP 和利率 3 标准偏离差的冲击；类似的，Castren、Fitzpatrick 和 Sydow 采用 GVAR 模型一个宏观经济变量 5 标准离差的冲击。

Thomas Breuer et al（2009）提出了另一种思路。他们用风险因子的马氏距离（Mahalanobis distance）来界定压力情景的可能性。Breuer 等人的思路是这样的：假定

一贷款资产组合，其每一头寸的价值取决于 n 个系统性风险因子， $r = (r_1, \dots, r_n)$ ，和 m 个异质风险因子， $\epsilon_1, \dots, \epsilon_m$ 。假定系统性风险因子 r 的服从椭圆分布（elliptical distribution）。风险因子 r 的分布的协方差矩阵和均值分别记为， Cov 和 μ 。异质风险因子的分布则比较主观。假定风险因子偏离其均值越远，其发生的可能性就越低。这样用标准离差测度的距离其本身也有概率含义。于是作者用马氏距离来界定情景的可能性：

$$Maha(r) = \sqrt{(r - \mu)^T Cov^{-1}(r - \mu)}$$

$Maha(r)$ 就代表了情景的可能性：马氏距离越大，意味着情景 r 的可能性越低。这种方法的优点主要有两个。一是它可以使风险管理者在情景的可能性和严峻性之间做权衡；二是它能够解决量纲依赖问题（dimensional dependence）²。

Klaus Dußmann, Martin Erdelmeier (2009) 认为，传统压力测试一个缺陷是集中于一个单一情景，这个情景发生的可能性很低。他们构建的情景由一系列压力事件组成，这样压力情景发生的可能性就比较高。他们在对 28 家德国银行信贷资产组合进行压力测试时，构建的情景发生的可能性高达 33%。

(2) 模型的选取

就基本框架而言，不管是 FSAP 国家还是非 FSAP 国家使用的框架都是一样的（见图）。区别主要在于模型的选取、变量的选取、数据的使用、估计方法和影响的测度上。在图中可以看出，我们需要两类模型来进行压力测试：一类是宏观经济模型，用以确定压力事件对宏观经济变量的影响；另一类则是风险模型，用以确定宏观经济变量对银行风险的影响。另外，由于本文内容的需要，在这里我们只讨论信用风险模型的选取。

(3) 宏观经济模型

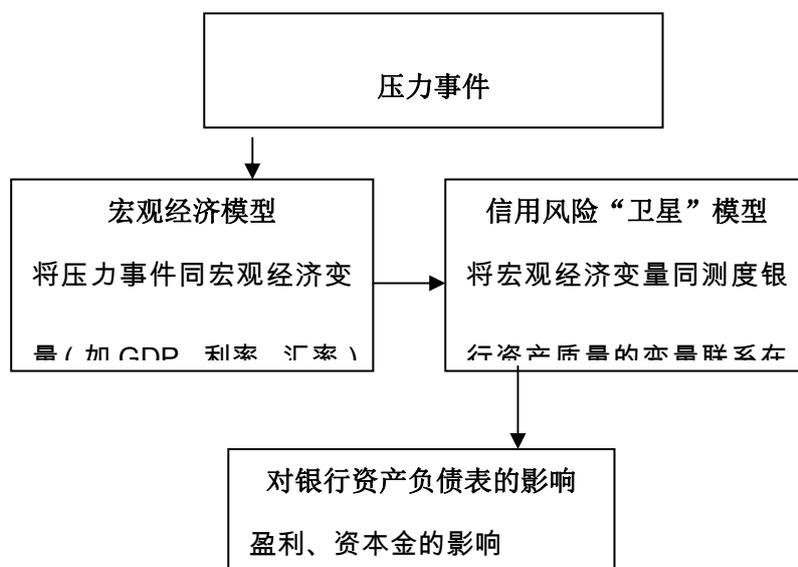
常用的宏观经济模型有三种：（1）结构性宏观经济模型；（2）向量自回归模型（vector autoregressive, VAR）；（3）纯统计模型。

FSAP 国家在进行宏观压力测试时通常使用的是国内货币政策分析的宏观经济模型，当冲击源是其他国家并通过国际联系传递进来时，这些模型不能提供相关信息。

Jones、Hilbers 和 Slack (2004) 对宏观结构模型的使用进行了探讨，包括基准假设的选择、政策反应、时间期限、那些变量保持固定、那些变量假设受到冲击。另一个经常提到的问题是，压力时期宏观经济变量相关关系会变得非线性，而这些线性模型不能反应这些特征。

图 1.3 信用风险压力测试传导图

²最大损失有一个奇怪的尺度依赖问题：对一个固定的资产组合和固定概率的可行域（admissibility domain），增加不相关的风险因子将会增大最大损失。Studer (1997) 曾举例论述此问题。假设一个债券资产组合，影响其价值的因素是 20 条利率曲线和 10 个国家货币。A 风险管理者对每条利率曲线用 7 个 buckets 建模，B 风险管理者用 15 个。这样 A 风险管理者用 150 个风险因子，B 用 310 个风险因子。两者都在 $a=95\%$ 的椭圆分布计算最大损失。 $n=150$ 时，A 得到马氏距离 $k=13.40$ 。B 用 $n=310$ ，将会得到 $k=18.76$ 。对线性资产组合，最大损失对马氏距离 k 的椭圆依赖是 k 的线性函数。因此，B 计算的最大损失将是 A 的 $18.76/13.40=1.4$ 倍。问题是，两个风险管理者采用的是相同的资产组合和相同的置信度 a 。



若缺乏合适的宏观经济模型，可采用向量自回归模型或向量误差校正模型（vector error correction model, VECM）。在这些模型中，初始冲击对一组宏观经济变量同时产生影响。VAR 模型的吸引力在于其灵活和相当简单的方式。其使用国家有英国（英格兰银行，BoE）、日本（日本银行，BoJ）、西班牙（西班牙央行，BoS）、荷兰（荷兰央行，DNB）以及欧洲央行（ECB）。

日本银行在其 2007 年的金融体系报告中（Financial System Report）估计了一个由 5 个宏观经济变量组成的 VAR 模型，其宏观经济变量包括 GDP、通货膨胀率、银行未清偿贷款、有效汇率和隔夜拆解利率）。Van den End, Hoeberichts, and Tabbae（2006， DNB）和 Jiménez and Mencía（2007， BoS）使用 VAR 模型对包含两个宏观经济因子的冲击反应进行建模。

Castrén, Dées 和 Zaher（2008， ECB）使用的模型是一个全球向量自回归模型（Global Vector Autoregressive, GVAR），基于国别或地区的误差校正模型 VECM，在这里国内和国外变量同时相互影响，在国别模型中，内生变量包括实际产出、通胀率、实际股权价格、短期和长期利率。

英格兰银行的样本模型使用的是一个两国版本的 GVAR 方法，仅仅对美国和英国经济进行建模，其宏观经济变量同 ECB 相同。在 BoE 方法中，GVAR 由两个 VAR 模型组成，一个用于英国，另一个用于美国。英国被当作一个开放的小国，美国代表剩下的国家。Haldane, Hall, Pezzini（2007）对 BoE 方法进行了讨论。BoE 方法对收入、市场风险和信用风险以及银行间的一些重要的反馈效应（如网络效应、市场流动性等外部性）共同建模。Alessandri, Gai, Kapadia, Mora 和 Puhr（2007， BoE）对这个模型进行了描述。

奥地利央行的 SRM 模型，用于金融体系稳定性分析和压力测试，其框架包括信用风险、市场风险、银行间传染风险（Boss, Breuer, Elsinger, Jandacka, Krenn, , Lehar, Puhr, and Summer, 2006）。

同结构性宏观经济模型和 VAR-VECM 模型相反，奥地利国民银行（Oesterreichische Nationalbank, OeNB）在其系统性风险监测（Systemic Risk Monitor, SRM）中发展了一个纯粹的统计方法来设计情景。宏观经济和金融变量通过一个多元 t 联系物（t-copula）建模。联系物方法有两个优势：一是边际分布同以变量联合行为为特征的多元分布不同。二是宏观经济变量和金融变量间的相互依赖（co-dependence）显示出尾部依赖（tail-dependence）情况（压力情景下相关关系增加）。但是作为一个纯粹的统计方法，这种方法不适合政策分

析，因为冲击的传导机制很难解释。

(4) 信用风险模型

结构宏观经济方法和 VAR 方法都需要一个模型把宏观经济变量映射到压力情景下银行资产负债表变动的指标上。典型的宏观经济变量并不包含信用风险测度，因此压力测试的第二步一般是对连接信用风险测度与宏观经济变量的卫星模型或辅助模型进行估计，将外部冲击映射到银行资产负债表之中。

在这些信贷质量回归模型中，贷款表现的测度典型地同宏观经济条件的测度相关。Blaschke、Jones、Majnoni、Martinez Peria (2001) 用不良贷款比率 (NPL) 对名义利率、通货膨胀率、实际 GDP 变动以及贸易变动进行回归分析。回归系数提供了贷款表现 (loan performance) 对这些宏观经济因子的敏感性估计。这种方法假设贷款质量同经济周期是敏感的。估计方法通常要求根据理论和实证选取的宏观经济变量和金融变量对信用风险有显著影响。Segoviano (2006) 详细分析了如何选取变量。

在卫星模型中，宏观经济变量是外生的，因此它忽略了银行体系对宏观经济的反馈效应，这也是传统压力测试的主要缺陷之一。Castrén、Dèes 和 Zaher (2008, ECB) 认为几个可能的原因是，缺乏足够长的时间序列数据，更灵活的建模方法，使压力测试实施起来更简单，结论也更易解释。

与宏观经济模型不同，信用风险卫星模型可以用个体银行甚至个体借款人的数据来估计。Cihák (2007) 将这种方法分为两类：一类基于贷款表现的数据，如不良贷款 (nonperforming loans, NPLs)、贷款损失准备金 (loan loss provisions, LLPs) 以及历史违约率；另一类基于同家庭部门或公司部门违约风险相关的微观水平的数据。很多信用风险卫星模型是对信贷资产组合损失分布的估计，这就对整体风险情况进行估计并对冲击的影响进行评估。

A. 基于贷款表现的模型

这种方法中，关键的应变量是 NPLs、LLPs 和历史违约频率。这些模型包括各种宏观经济因子，不同的国家选取的变量有所不同 (表 1)。有时，一些模型会添加同企业资信状况直接相关的变量，如债务；在另一些情况，则可能会添加市场基础的信用风险指标，如股票价格、公司债券幅差。

Alessandri、Gai、Kapadia、Mora、Puhr (2007, BoE) 和 Marcucci、Quagliariello (2005, BoI) 使用家庭或公司部门违约频率对信贷质量进行建模。Marcucci 和 Quagliariello 使用 VAR 方法来估计信用风险卫星模型，他们对公司部门的建模包括违约率和四个宏观经济变量 (产出缺口、通货膨胀率、短期利率、实际汇率)。违约率假定对产出缺口和其他宏观经济变量是外生的。冲击响应函数 (impulse response function, 也称脉冲反应函数) 表明各种宏观经济变量 (通胀率除外) 对违约率有显著影响。

德国央行、瑞士央行和荷兰央行 (Lehmann, Manz, 2006; Van den End, Hoeberichts, and Tabbae, 2006) 的信用风险模型对 LLPs 使用静态和动态的面板数据估计来测度个体银行水平的信贷质量。对个体银行 LLPs 的面板估计控制影响信用风险的个体银行特征，并获取银行对宏观经济发展的不同敏感性。

Fiori、Foglia 和 Iannotti (2008, BoI), Jiménez 和 Mencía (2007, BoS) 以及奥地利央行的 SRM 系统都使用按照产业分组的历史违约率建模。这种部门统计分析可以使用不同的宏观经济变量来解释不同行业的违约频率，并可添加具体部门的解释变量以提高拟合度。

在这些模型中，对很多部门都有显著影响的宏观经济变量代表系统性风险成分；部门间违约相关性是由于对系统性风险成分的共同依赖。异质风险由潜在的具体部门变量和部门方

程残差来测度。

Fiori、Foglia、Iannotti (2008, BoI) 和 Jiménez、Mencia (2007, BoS) 认为部门间的微观传染效应是违约相关性的一个额外渠道。这两篇文章允许部门违约频率受宏观经济条件和能反应传染效应的潜在因子的影响。相应的，他们将“周期性”部门（对系统性风险高度敏感）和受异质风险影响大的部门区分开。两篇文章都发现显著的微观传染效应，并将农业、制造业、建筑业、贸易行业看做周期性部门，矿产业、公共事业看做异质性部门。

使用贷款表现数据来测度信贷质量有一些问题。贷款表现是一个滞后的、回顾性的指标，反应的是过去的违约。贷款损失准备准则在不同地区可能有所不同：不同的法律准则可能导致机构在注销不良贷款或提取准备金上有所不同；LLPs 的变化可能只有一部分是由信用风险的变化引起的；特定银行的因素，如收入熨平政策 (income-smoothing policies)，可能也会发挥作用。

另一个问题是违约率的非线性转换问题。关于信用风险的宏观经济模型通常使用线性的统计模型。在冲击很小的时候，线性模型是可以的，但是当冲击很大时，非线性模型会更加重要。Wilson (1997) 曾对这方面问题进行综述，并使用非线性转换，如 logit 和 probit 转换，对违约率进行建模。Van den End、Hoeberichts 和 Tabbae (2006, DNB) 认为，违约率的非线性转换将解释变量的范围扩展到负值，并考虑了压力情景下宏观经济变量和违约率可能的非线性关系。

B. 基于个体借款人数据的模型

这种方法使用个体借款人数据来估计信用风险卫星模型。在这种情况下，模型也可能使用宏观金融数据。如果模型没有包含宏观经济变量，那可能会使用一个额外的卫星模型把宏观金融变量与具体借款人数据联系起来。

Eklund、Larsen 和 Berhardsen (2001, Norges Bank) 使用挪威的所有有限责任公司的年度会计数据，将违约概率与借款人特征（如企业年限、规模、所属行业）和测度公司盈利、流动性以及财务能力的会计变量联系在一起。主要宏观经济变量的预期值被用于估计每家公司未来的资产负债表和损益表情况，在此基础上计算个体违约概率 (PDs)。用违约概率就可以得到银行部门贷款总损失的估计值。

个体信贷质量测度可以被开发来估计其与宏观经济变量的直接关系。Asberg、Shahnazarian (2008, Sveriges Riksbank) 和 Castren、Fitzpatrick、Sydow (2008, ECB) 使用穆迪的 KMV EDFs 来对上市公司平均信贷质量建模。EDF 是信用风险的前瞻性的、市场基础的测度，它基于一家公司股票波动率来估计其一年内的违约可能性。

在 Asberg 和 Shahnazarian (2008) 的文章中，瑞典所有非金融上市公司的 EDF 中位数作为违约可能性的替代。作者使用 VECM 对总的 EDF 与三个宏观经济变量（行业产出指数、消费者价格指数、短期利率）进行估计。假设变量之间存在长期相关性，VECM 能发现序列和短期波动间的共同趋势。研究结论表明，对 EDF 有强烈的、正的影响的宏观经济变量是利率；制造业产出的降低和通胀率的上升同样导致更高的 EDF。Asberg 和 Shahnazarian (2008) 发现，较高的通胀率意味着更高的因子价格，这会导致成本的增加并倾向于影响信贷质量。而且，高通胀通常被认为是宏观经济管理失当的信号和不确定性的来源。因此，违约率和通胀率之间的关系应该是正的。但是，较高的通胀率也意味着更高的产品价格，这会促进收入的增加，降低债务负担，因此改善资信状况。

Castren、Fitzpatrick 和 Sydow (2008, ECB) 用欧元区公司 EDF 的中位数度量信用风险。模型将公司的信贷质量同 5 个宏观经济变量以及股票价格联系在一起。其研究表明，利率对信贷质量的影响并不显著。这个结论有点违背直觉，因为通常认为利率对公司信贷质量有很重要的影响。作者用应变量的特征来解释：EDFs 的主要驱动因素是市场资本值

(market capitalization) 和违约点 (它是债务的函数)。

与使用市场基础的信用风险测度相反, 法国银行委员会 (FBC) 和日本央行使用个体非金融企业信用评级的数据来进行分析 (Commission Bancaire, 2007; Bank of Japan, 2007)。这两个模型都估计借款人信用评级转移概率对宏观经济变量的敏感性。在 FBC 的模型中, 宏观经济变量是 GDP、短期利率和长期利率。考虑到误差项之间可能的相关关系, BoJ 模型用似不相关回归 (Seemingly Unrelated Regression) 估计了 5 个方程 (每一评级一个方程), 其解释变量是 GDP 增长率和一个杠杆率 (作为盈利和债务状况的替代)。

(4) 压力测试的实施

这一步要做的是用宏观经济模型来预测压力条件下宏观经济变量的数值, 然后用信用风险模型来估计银行在压力条件下的信贷质量。

FSAP 框架使用的是宏观计量经济结构模型。这种方法下, 模型生成的点估计值有一个单一的未来路径, 这个路径是压力情景下的条件平均路径, 没有概率分析。VAR/VECM 框架能够生成具有概率分析的压力情景。冲击的规模由自回归序列中无条件标准离差确定, 在正态分布假设下, 冲击的规模就会对应一个概率。这类模拟的尾部就代表极端情景。这类情景同宏观结构模型的区别在于, 它仅仅基于一个概率方法, 没有经济学理论支撑。

Pesaran、Schuermann、Treutler 和 Weiner (2006) 最早使用 VAR 模型生成一个概率情景来分析信用风险。作者使用冲击响应函数来检测对一个宏观经济变量单独的冲击如何影响其他变量。冲击反应函数假定其他变量根据它们同被冲击变量的历史协方差被替换, 因此这种方法恰当的考虑了冲击间的相互关系。这种方法对特定问题风险管理特别有价值, 特别的, 它可以产生一个可能的冲击情景等级序列。

Jiménez、Mencía (2007, BoS) 和 Castren、Fitzpatrick、Sydow (2008, ECB) 以及日本银行在压力测试中采用了这种概率方法生成情景。在 BoJ 的模型中, 压力测试估计一个负的 GDP 冲击的影响, 冲击的大小对应的概率为 1%。Jiménez 和 Mencía 采用一个 GDP 和利率 3 标准偏离差的冲击; 类似的, Castren、Fitzpatrick 和 Sydow 采用 GVAR 模型一个宏观经济变量 5 标准离差的冲击。

奥地利的 SRM 多元 t-copula 方法被用来描述风险因子的随机变化。根据既定冲击, 一个或多个模拟的风险因子的变化被设定一个固定值, 其他风险因子 (没施加压力的) 的变化由压力情景的条件分布描述 (Boss, Breuer, Elsinger, Jandacka, Krenn, Lehar, Pühr, Summer, 2006)。

Van den End、Hoeberichts、Tabbae (2006, DNB) 提出了另一种方法, 这个方法考虑了宏观经济变量的同时变化和它们之间的相互作用。宏观经济变量预期值被用于一个 VAR 模型以重估 GDP 和利率。VAR 重估考虑了相关关系的变化, 这样可以克服这个问题——压力测试模型所假定的相关关系在压力情景下可能并不正确的问题。Asberg、Shahnazarian (2008, Sveriges Riksbank) 采用了类似的步骤。对于给定冲击, 他们使用了 Riksbank 宏观计量经济模型的冲击响应函数来估计三个宏观经济变量的“压力值”, 然后用 VEC 模型来预测压力 EDFs。

(5) 影响测度

最后一个步骤是评估压力事件对银行贷款的影响, 并判断银行是否能够抵抗假定的冲击。这意味着将损失和一个合适的基准进行比较。按照所使用的信用风险模型, 模拟的结论可以用预期违约率或损失准备金来表达。基准情景, 银行是盈利的, 所以压力情景下的损失首先冲减银行利润, 然后才是资本。若压力情景下的损失直接表述为资本的话, 无疑会高估压力事件的影响。

影响测度的一个重要扩展是试图得到银行体系贷款的损失分布, 而以往的方法是产生一个压力事件的会计测度——压力情景下的点估计。在的款损失分布的情况下, 压力测试可以

用参数的移动来表述，如 PD、LGD (loss-given-default)。宏观压力情景可以用来模拟不利的宏观经济条件，生成一个总的“压力”PD 或一组部门 PDs。通过这种连接，压力测试就有一个清晰的经济分析。

Sorge 和 Virolainen (2005) 最先提出用银行体系的信贷损失分布来测度信贷冲击的影响。Pesaran、Schuermann、Treutler 和 Weiner (2006) 有类似的研究。很多国家的监管当局对这一思路进行了完善。一种方法是使用蒙特卡洛模拟技术来估计资产组合的损失分布 (Alessandri, Gai, Kapadia, Mora, Pühr, 2007, BoE; Jiménez, Mencia, 2007, BoS)。第二种方法是使用现有的资产组合模型，如 Credit Risk Plus，来模拟宏观经济因子的随机变动 (Castren, Fitzpatrick, Sydow, 2008, ECB; Sveriges Riksbank, 2006)。

二、国内压力测试研究现状

国内对压力测试的研究也分为两个阶段，早期研究主要集中在介绍国内压力测试方法，近几年国内学者也开始结合我国实际探索适合我国的压力测试方法、模型。大部分的研究集中于宏观压力测试和信用风险压力测试两个方面。

(一) 方法、框架类

唐文江等 (2009) 对压力测试情景设置进行了探讨，分析了情景设置的方法，并对情景设置中如何建立宏观经济因子之间的联系、风险因子的传导机制以及情景设置如何与风险管理文化结合、应对突发事件等问题进行了探讨。

巴曙松、朱元倩 (2010) 总结了压力测试在国际上的实践规范、压力测试方法，以及压力测试执行过程中的频率、时间跨度、执行流程等问题，对压力测试中的热点问题如事件冲击到承压变量之间的传导机制、缺乏数据情况下的压力测试以及宏观压力测试等问题进行了探讨。

徐明东、刘晓星 (2008) 深入研究了宏观压力测试的理论模型和执行宏观压力测试的主要步骤、方法，并对宏观压力测试的主要难题如银行之间的相互影响与反馈效应等进行了研究。他们的分析主要基于对 Goodhart 的违约依赖模型、Boss 的网络模型以及 Tsatsaronis 的信用风险转移模型几个模型比较之上。同时，他们也针对目前主要的宏观压力测试系统，包括 FSAP 系统、奥地利央行的 SRM 系统以及英格兰银行的 TD 系统，进行了比较研究。

刘晓星 (2009) 对压力测试和风险价值这两种风险管理技术进行了理论探讨，分析了压力测试在金融体系稳定评估方面的作用和实施步骤，并就压力测试中常见的反馈效应、传染效应、交互效应等难题进行了探讨。

陈颖、李婧 (2009) 从监管部门和商业银行两方面分析了在 FSAP 框架下开展宏观经济压力测试的相关问题。对监管当局而言，如何选取适合自己国家的压力测试情景，如何对情景进行建模以反映情景情况下宏观经济变量的变化，以及如何对压力测试结论进行解释是其面临的主要挑战；对商业银行而言，如何准确判断自身面临的主要风险并据此构造情景，如何协调各部门进行压力测试并根据压力测试结果做出反应，是其面临的主要问题。陈颖、李婧 (2009) 对这些相关问题提出了解决办法。

黄志凌 (2010) 对银行如何开展压力测试进行了较为深入的研究，提出了很多有价值的见解。他指出：(1) 压力测试要同日常风险管理工具区分开，要关注于极端情景下的风险，它与传统风险管理工具在使用范围和内在逻辑上是有本质差异的；(2) 尽管数据收集很难，但压力测试还是要尽量用数据说话，避免受主观因素或风险偏好的影响；(3) 不同银行有不同的客户群体和资产结构，这决定了它们对极端情景的风险敏感性存在差异，因此压力测试的选择上也会不同。(4) 压力测试模型要充分反映“压力下的变化”。

杨晓奇 (2010) 认为我国经济运行情况同发达国家有很多差异，国外的很多理论模型不太适合我国。因此杨晓奇试图通过寻找数据变化规律来分析系统性风险。他选取不良贷款率为指标评估银行体系稳定性，在 Wilson (1997) 的研究框架下，使用蒙特卡洛模拟和误差

校正模型进行宏观压力测试，评估银行体系在 GDP 增长率为 7%、6%、5%情景下银行体系的稳定性。同时杨晓奇在压力测试中还考虑了政府针对压力情景可能采取的宏观调控政策效果。其研究结论表明，随着压力程度的增大，不良贷款率也逐步上升，但政府调控政策对增强稳定性效果显著。

徐光林（2008）采用了一个新的角度来研究银行体系稳定性。他选取物价水平和 GDP 增长速度为压力指标，同时选取贷款利率为辅助指标。压力情景设置时参照了近 30 年我国国内生产总值增长速度和物价水平增长速度，同时结合测试时宏观经济环境，采用 5%的 GDP 增长速度和 14%的通货膨胀率为重度情景。同时，作者根据对 GDP 和物价水平长期、短期相关性分析，指出短期内 CPI 与 GDP 增长速度相关性低，进而假定 CPI 冲击量的确定与 GDP 冲击量的确定相互独立。同时，作者还认为我国利率水平虽然受政府决定，但政府的利率政策会受各种市场力量的影响，因此利率变动与 GDP 和 CPI 密切相关。同时作者假定风险因子变动服从指数分布。在此基础上，作者采用线性模型，分析 GDP、CPI 以及利率的不利变动对银行资产规模的影响。值得注意的是作者还采用了 Brumer（2008）的方法，利用马氏距离来判断压力情景发生的可能性。

（二）信用风险类

国内学者对信用风险压力测试的研究主要分为两类。一类集中于整个银行体系信用风险研究，这类研究大多采用 Wilson（1997）框架，对宏观经济变量与信用风险指标直接进行统计计量分析，包括华晓龙（2009）、沈阳、冯望舒（2010）、周源（2010）、吴婷和段明明（2009）

等。另一类则主要集中在银行房地产贷款压力测试，分析房地产行业对金融机构和银行体系稳定性的影响，包括冯佳、朱华彬（2009）、陈韬（2009）、陈阳、陈双杰（2009）等。

华晓龙（2009）采用压力测试方法评估我国商业银行面临的信用风险。他的研究构建了两个压力情景，一是通货膨胀率大幅上升，二是 GDP 大幅下降。其选取的信用风险度量指标为贷款违约率，然后通过 logit 转换，将贷款违约率与宏观经济因子直接进行回归分析。其研究结论表明，在压力情景下，银行体系贷款违约率显著提高，银行体系风险大幅上升，而名义贷款利率、房地产价格指数、消费者价格指数、名义 GDP 的变化与贷款违约率密切相关。

沈阳、冯望舒（2010）借鉴 Wilson（1997）、Boss（2002）以及 Virolanien（2004）的模型，构建宏观经济变量与信用风险的联立方程组。他们选取不良贷款率为信用风险水平度量指标，在国内生产总值大幅下降和贷款利率大幅上升两个情景下，分析宏观经济指标变动对信用风险的影响。其结果表明，我国商业银行信用风险与主要宏观经济变量之间关系密切，消费者物价指数、房地产价格指数、贷款利率、GDP 增长率以及进出口总额增长率对不良贷款率影响显著。而相对于国内生产总值大幅下滑，贷款利率的大幅上升对银行不良贷款率的影响更显著。

陈阳、陈双杰（2009）认为依靠传统信用风险模型来执行的压力测试无法准确反应风险因子在压力环境下的变化，因此他们采用随机模拟方法来估算房地产开发行业在压力情景下的违约概率。其构建的压力情景为房价大幅下跌，中间冲击变量则为贷款利率和房屋销售面积。他们认为，无论企业违约的原因是什么，企业违约前都会表现出资金紧张状况，因此分析房地产开发企业的现金流，分析其资金缺口的期限、性质和大小就能有效识别房地产商的违约行为。陈阳、陈双杰（2009）采用蒙特卡洛模拟法，通过分析压力情景下，企业销售收入和融资成本的变动，进而分析企业现金流的变动。而随机模拟的现金流为负的频率则代表了房地产开发商违约的概率。在他们的方法中，房价和利率冲击都是随机的，因而现金流也是随机的。作者利用上市房地产开发商的数据，随机模拟了 8000 个情景，其研究结论表明房价下降 15%，房地产开发商违约概率大幅上升。

冯佳、朱华彬（2009）采用敏感性方法分析房价下降对开发贷款违约率和个人住房贷款违约率影响。他们结合我国 A 股市场 14 家上市银行的具体情况，利用压力测试测算房价下跌时房地产贷款违约对银行业净利润的影响程度，进而判断银行房地产贷款的风险。

三、国内外研究评述

早期的压力测试主要集中在压力测试的执行框架、操作流程上，近期的压力测试则从情景的设置、模型的选取、压力测试的执行、影响测度等多个角度出发对压力测试进行了较为深入的研究。目前的研究呈现以下几个特征：（1）大部分的研究由各国监管当局主导，研究的类型主要集中在宏观压力测试领域。（2）如何选取一个合适的情景，既保证压力充分又具备发生的可能性并在尽可能给出发生的概率，仍然是一个重点关注的问题。（3）在模型的选取上，如何确保模型假设条件在极端环境下仍然适用，并充分考虑风险因子之间的相互关系和反馈效应，是压力测试面临的又一难题。虽然一些学者做出了一些有益的探讨，但效果尚不明显。（4）对于影响的测度，主流的做法试图寻找压力情景下资产组合损失的分布，这种方法对数据和统计方法的要求较高，操作难度较大。

参考文献：

1. Andersen, H., T. O. Berge, E. Bernhardsen, K.-G. Lindquist, and B.H. Vatne. 2008. "A Suite-of-Models Approach to Stress-Testing Financial Stability." Staff Memo 2008/2, Norges Bank.
2. Alessandri, Gai, Kapadia, Mora and Puhr (2007), A framework for quantifying systemic stability, December, preliminary paper presented at the workshop "Stress Testing of Credit Risk Portfolios: The Link between Macro and Micro", hosted by the BCBS and the DNB, Amsterdam, 7 March 2008.
3. Bank of Japan (2007), The Framework for Macro Stress-Testing of Credit Risk: Incorporating Transition in Borrower Classifications, Financial System Report, September.
4. Berge and Lindquist (2007), A suite-of-models approach to stress testing financial stability, preliminary paper presented at the "2nd Expert Forum on Advanced Techniques on Stress Testing: Applications for Supervisors", IMF and DNB, Amsterdam, October 23-24.
5. Blashke, Jones, Majnoni, and Martinez Peria (2001), Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences, IMF WP, no. 88.
6. Bonti, Kalkbrener, Lotz, and Stahl (2006), Credit risk concentration under stress, Journal of Credit Risk, vo.2, no.3, 115-136.
- BOSS (2003), A macroeconomic model for stress testing the Austrian credit portfolio, OeNB Financial Stability Report, Vol. 4, pp. 64-82.
7. Boss, Breuer, Elinsger, Jandacka, Kpenn, Lehar, Puhr, and Summer (2006), Systemic Risk Monitor: A Model for Systemic Risk Analysis and Stress Testing of Banking Systems, OeNB Internal technical report.
8. 唐文江等. 国际银行业新资本协议压力测试情景设置探讨及对我国的借鉴. 国际金融, 2009
9. 巴曙松, 朱元倩. 压力测试在风险管理中的应用. 经济学家, 2010
10. 刘晓星. 风险价值、压力测试与金融系统稳定性评估. 财经问题研究, 2009
11. 徐明东, 刘晓星. 金融系统稳定性评估: 基于宏观压力测试方法的国际比较. 国际金融研究, 2008
12. 华晓龙. 基于宏观压力测试方法的商业银行体系信用风险评估. 数量经济技术经济研究, 2009
13. 沈阳, 冯望舒. 宏观经济变量与银行信用风险的实证研究——基于宏观压力测试的分析. 会计之友, 2010 (8)
14. 周源. 宏观经济数据影响下的信用风险压力测试研究. 金融纵横, 2010
15. 吴婷、段明明. 宏观经济因素与商业银行信用风险——实证与压力测试. 金融观察, 2009
16. 徐光林. 我国银行业金融机构资产规模的宏观压力测试. 新金融, 2008

17. 杨晓奇. 后危机时代宏观压力测试研究. 新金融, 2010
18. 陈阳, 陈双杰. 房地产开发企业违约概率压力测试研究——现金流蒙特卡洛模拟方法在银行中的应用. 金融论坛, 2009
19. 陈韬. 房价压力测试方法探讨. 金融管理与研究, 2009 (2)
20. 黄志凌. 商业银行压力测试. 北京: 中国金融出版社, 2010

Review of research on commercial bank stress tests at home and abroad

Peng Zhihui

(Chinese Financial Research Centre of Southwestern University of Finance and
Economics, Chengdu, 610074)

Abstract: Stress testing is a management technique which aims at testing banks' compressive capacity in extreme situation. This paper embarks from the foreign pressure test research situation, analyses the early stress testing framework, as well as the pressure test of the progress of the technology; Secondly, this paper expounds two kinds of domestic research: macro stress testing and credit risk testing; The last reviewed the research status at home and abroad.

Key words: stress tests commercial bank risk management credit risk

收稿日期: 2013-1-15

作者简介: 彭志慧, 西南财经大学中国金融研究中心金融学博士生, 研究方向: 金融理论与实践