Stress Test Robustness: Recent Advances and Open Problems

Thomas Breuer,
Martin Summer[1]

This paper reviews recent advances made in improving the robustness of stress-testing models against potential misspecification or risk-factor-distribution misestimation, including conceptual advances in measuring robustness against pricing-model misspecification. In addition, we address an important open problem of stress tests as they are carried out today: the endogeneity of financial risks. Traditional stress-testing frameworks model a single-person decision problem in the face of an exogenous source of risk. Yet financial risks arise from the complex interaction between individuals, firms and financial institutions. A stress-testing framework that falls short of incorporating this risk endogeneity will ultimately only be able to capture the financial stress of individual institutions in a non-crisis environment.

JEL classification: G01, G28, G38, C02
Keywords: stress testing, financial stability, systemic risk, robustness

1 Introduction

The idea of stress testing financial portfolios stems from the realm of risk management.[2] Risk managers use stress tests to identify possible scenarios that would be extremely damaging to the value of the current portfolio, and to quantify the losses that might occur under such detrimental scenarios. Stress tests are meant to help financial institutions figure out whether their ultimate risk-bearing capacity is sufficient to remain solvent even in an extremely difficult economic environment. Lately, even entire financial systems have been subjected to "macroprudential" stress tests. While stress tests used to receive attention only in small circles of risk management professionals and regulators, they have gained broader public attention during the recent financial crisis. The U.S. Federal Reserve System has been mandated to perform annual stress tests of major financial institutions under the Dodd-Frank Act, and also the newly created European Banking Authority will conduct stress tests for European banks on a regular basis. Despite this policy prominence, the methodology of stress testing is still in its infancy and needs further development. In this paper we discuss some recent advances that might improve stress testing and identify some open issues. We provide an overview of recent research evidence on how to make stress tests more robust against model misspecification within the traditional stress-testing framework. Yet while more robustness is desirable we believe that, ultimately, the stress-testing framework as such needs to be enhanced to capture the economic nature of financial crisis more adequately. In this respect we also discuss what we consider to be the most important open problem of current stress tests.

2 Statistical Risk Models

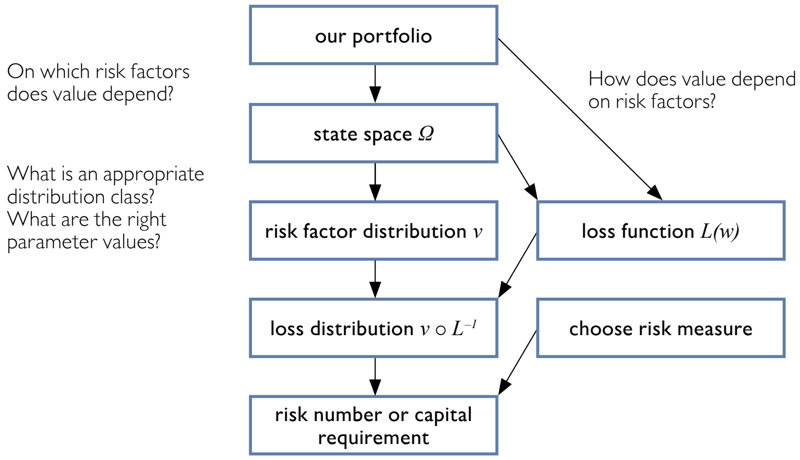
Most of the current stress-testing methods are based on concepts derived from statistical risk models, with a risk factor distribution serving as the foundation of the abstract framework. Within this framework, the individual risk factors are coordinates of the state space Ω, which codifies our lack of

[1] *Fachhochschule Vorarlberg/PPE Research Center, thomas.breuer@fhv.at and Oesterreichische Nationalbank, Economic Studies Division, martin.summer@oenb.at.*

[2] *For a standard textbook treatment of stress testing, see Jorion (2000). An important supervisory reference is the Basel Committee on Banking Supervision (2009).*

Chart 1

Standard Risk Measurement Procedure



Source: Authors' analysis.

knowledge regarding all uncertain events that may affect the value of a given portfolio, financial institution or system of financial institutions. The resulting risk factor distribution v assigns a probability to every event. It summarizes our statistical knowledge about the system in question. Often it results from the choice of a model class, or from a parameter estimation procedure based on historical data.

Within this framework, the financial portfolios that the statistical risk models are meant to assess are described by real functions L on the state space. The random variable $L(\omega)$ can be thought of as the disutility of the portfolio outcome described by L if state ω occurs. We refer to L as the loss-pricing function.

The basic structure of a standard portfolio stress-testing model is illustrated by chart 1: The state space Ω is determined by the risk factors on which the value of the portfolio depends. In turn, the specification of a distribution class and a parameter estimation procedure applied to historical data determine the risk factor distribution v. How the portfolio value depends on the risk factors is described by the loss-pricing function L. Each of these three steps involves modeling decisions and is hence a potential source of misspecification. Both the risk factor distribution and the loss function determine the distribution of losses $v \circ L^{-1}$. The model thus produces a risk measure that assigns a risk number or a capital requirement to the loss distribution.

Based on this structure, we distinguish between two main types of model risk. *Distribution model risk* stems from statistical model misspecifications or from parameter estimation errors, and leads to a wrong risk factor distribution v. *Pricing model risk* stems from modeling errors concerning the dependence of the portfolio value on the risk factor values. It leads to a wrong loss-pricing function L.

Example 1. *For a linear portfolio the loss is given by a linear function of n risk factors. In vector notation the loss function is* $L(\omega)=l\cdot(\mu-\omega)$*, where vector* ω *of the risk factor values is modeled as normally distributed with mean* μ *and covariance matrix* Σ*,* $\omega \sim v = N(\mu,\Sigma)$*. The vector* l *describes the portfolio weights. (For equity portfolios, the specification must be adjusted as the value of*

Table 1

Estimating Rating Class Transitions and Related Losses (Example)

	AA1–2	AA3	A	BBB	BB	Default
	%					
Loss from rating class transition	–3.20	–1.07	0.00	3.75	15.83	51.80
Estimated transition probabilities	0.09	2.60	90.75	5.50	1.00	0.06

Source: Authors' analysis.

a stock cannot fall below zero. The solution is to either assume the risk factors to be distributed log-normally rather than normally, or to take the risk factors to be log-returns, in which case the loss function is exponential rather than linear.)

Example 2. *A simple credit risk model can be used to assess the n states in which an obligor may find himself or herself at some future time, i.e. to indicate the probabilities of a transition from the current rating class to some rating i by* p_i*. The model serves to estimate the reference risk factor distribution v based on historical data and to produce a traditional transition matrix where each column represents a vector* $v=(p_1,\ldots,p_n)$*. For each possible final rating, the loss* l_i *caused by a transition into that class is specified by market data and obligor data.*

The table above provides a numerical example for an A-rated bond. Model estimations show the probabilities with which this rating can migrate into other rating classes (second row of the table) and the losses that are to be expected (first row). (These loss numbers were determined from credit spreads of A-rated industrial bonds maturing in five years, as given by Bloomberg.) Under the estimated transition probabilities the expected loss is 0.37% of the bond value.

3 Stress Tests as Scenario Analysis

In its early days, stress testing was intended to provide risk information about an institution or portfolio without relying on a specific risk factor distribution, which might be misspecified or misestimated. The approach was to evaluate a simple loss function L at certain scenarios. The scenarios themselves were chosen in an informal discussion among experts on potential risk factor realizations that are regarded extreme yet plausible. The precise meaning of these terms was left undefined. Sometimes existing models were used to construct the scenarios, like the central banks' macro forecasting model, sometimes scenarios were chosen based on historical experience or other considerations.

Example 3. *For the linear portfolio of example 1, one scenario could be a 3σ drop of all risk factors from their current values μ,* $\omega=\mu-3(\sqrt{\Sigma_{11}},\ldots,\sqrt{\Sigma_{nn}})$*. The resulting loss would be* $L(\omega)=3(l_1\sqrt{\Sigma_{11}}+\ldots+l_n\sqrt{\Sigma_{nn}})$*. For the credit risk model of example 2, one scenario could be a downgrade of the bond from A to BB, leading to a loss of 15.83%.*

If the loss resulting from the given scenarios $L(\omega)$ was deemed unacceptable, the institutions needed to decide whether the scenario ω was plausible enough to warrant counteraction, and to determine what this counteraction could be. This "scenario analysis" procedure is still popular and continues to underlie most of the stress tests currently performed. Financial institutions invest very substantial efforts to translate scenarios provided by supervisors in terms of a handful of macroeconomic risk factors into risk factor moves of the institution's internal risk model.

This approach suffers, however, from two important drawbacks. First, a stress test that comes up with acceptable results for the scenarios analyzed may provide an unjustified illusion of safety, as it does not provide any information about any other scenarios that were not taken into consideration. Banks may, after all, become insolvent despite having passed recent stress tests. A notable example are the supposedly successful stress tests of Irish banks in 2010, which had to be bailed out a few months later.

Second, the judgment whether an alarming stress-test result warrants counteraction is necessarily based on a concept of plausibility, be it explicit or implicit. If stress scenarios are highly implausible, an alarming stress-test result may trigger a false alarm. The plausibility concept of scenarios should somehow be based on information about risk factor distribution. However, if this concept requires exact knowledge of the distribution it threatens to undermine the original purpose of stress testing, namely to provide information about an institution without relying on a specific risk factor distribution.

A first attempt to overcome the two drawbacks was made by Studer (1997, 1999) and Breuer and Krenn (1999), who developed what one could call "traditional systematic stress tests." Their approach, used in the context of multivariate normal risk factor distributions, is to first select ellipsoids (of a specified Mahalanobis radius) to arrive at a set of sufficiently plausible scenarios and then to identify the worst case among those scenarios. This approach addresses the two drawbacks: It does not sound a false alarm because only scenarios of sufficient plausibility are considered, and it does not create a false illusion of safety because the worst-case search ensures that no dangerous and plausible scenarios are neglected. This approach is probably a sensible compromise between presupposing exact knowledge of the risk factor distribution and assuming complete ignorance about the distribution. It uses some distributional information in the definition of the set of plausible scenarios. All scenarios within this set are on the same footing; all scenarios outside this set are neglected. Hence the infinity of possible density values of scenarios is reduced to the two values "in" and "out."

Breuer et al. (2013) apply these ideas to a comparative stress-testing exercise for a big aggregate loan portfolio based on loan data from the Spanish loan register. They show that, compared to standard stress-test procedures, worst-case searches of plausible domains identify more harmful scenarios that are equally plausible than the scenarios considered in standard procedures.

While this approach solves the problems of creating false illusions of safety or false alarms, it has problems of its own. First, choosing a Mahalanobis ellipsoid as a scenario set is natural only for elliptical risk factor distributions, like normal or Student *t*-distributions. It is not clear how to choose sets of plausible scenarios if the risk factor distribution is not elliptical. For example, how should systematic stress tests be performed in credit risk models with discrete rating classes? Second, the stress-testing procedure is subject to model risk because it has to commit to a specific risk factor distribution. It is not robust with respect to the misspecification of risk factor distribution: The stress tester does not know by how much worst-case losses differ if the risk factor distribution is different from the one anticipated. Third, the Mahalanobis distance as a plausibility measure reflects only the first two moments of the risk factor distribution. This is not in line with intuition. A given extreme scenario should be more plausible if the risk factor distribution has fatter tails. Fourth, the maximum loss over a Mahalanobis ellipsoid depends on the choice of coordinates, as pointed out in Breuer (2008). Fifth, the worst-case loss over the ellipsoid is not a law-invariant risk measure: portfolios might have the same profit-loss distribution without having the same worst-case loss.

4 Systematic Stress Tests for Distribution Model Risk

In this section we focus on one important form of model risk, namely risk-factor-distribution uncertainty. Given the variety of opinions of economists and analysts about future average asset returns, correlations or volatilities, chances are that only one model, if at all, will correctly anticipate the risk factor distribution. All others, if not all, must bear the consequences of model risk. In the finance literature the term "model risk" frequently refers to uncertainty about the risk factor distribution, see Gibson (2000). We do not follow this convention, although the term is sometimes used in a wider sense (see Crouhy et al., 1998). As we take the term, it is equivalent to ambiguity in the sense of Ellsberg (1961).

The canonical expression describing worst-case expected losses over a set of alternative risk factor distributions is

$$\sup_{Q \in \Gamma} E_Q(L) \qquad (1)$$

for some closed convex set Γ of risk factor distributions. Problem (1) is a formal model of both stress testing and model risk. In the context of stress testing the distributions Q in expression (1) are interpreted as smeared versions of a particular scenario ω. In the context of model risk the distributions Q are plausible alternatives to the risk factor distribution ν.

Expression (1) provides a starting point for generalizing the idea of worst-case search over plausible domains. The key idea is to think in terms of sets of risk factor distributions rather than a particular risk factor distribution. The stress test is then carried out with a view to identifying the worst expected loss of the portfolio across risk factor distributions in the set Γ. This leaves the question of how Γ should be chosen. One key idea, advanced by Hansen and Sargent (2001), is to work with the set of all distributions whose relative entropy with respect to an initial prior distribution is below a given threshold. This leads to a generalized version of worst-case search over plausible domains where the threshold on relative entropy provides a plausibility constraint.

4.1 Relative Entropy as a Measure of Model Plausibility

We take the set Γ as the set of distributions which have relative entropy with respect to best-guess distribution ν smaller than some threshold $k>0$:

$$\Gamma = \{Q : D(Q \| \nu) \le k\}. \qquad (2)$$

The relative entropy $D(Q\|\nu)$ is taken as a measure of (im)plausibility of a particular risk factor distribution Q when the distribution-class specification and an estimation process lead to the best-guess distribution ν. This choice for the set Γ amounts to considering *all and only* the risk factor distributions above the plausibility level k. It encompasses both estimation risk and misspecification risk in the sense of Kerkhof et al. (2010).

In the literature, various "distances" of probability distributions are used.[3] One family of such distances – the *f*-divergences of Csiszár (1963), Ali and Silvey (1966) and Csiszár (1967) – corresponds to convex functions *f* on the positive numbers. Relative entropy corresponds to $f(t) = t \log t$, several other choices of *f* also give distances often used in statistics. Another important family are the Bregman (1967) distances, which also contain relative entropy as a special case. The results reported below can be generalized to sets Γ which are balls for *f*-divergences or for Bregman distances.

Relative entropy has already been used in econometrics; see Golan et al. (1996), Avellaneda and Paras (1996), Avellaneda et al. (1997) and Borwein et al. (2003). Using relative entropy in stress tests as well appears reasonable, though we do not claim that it is necessarily the best choice among the various distances of distributions. Relative entropy balls are a popular choice for describing model uncertainty in portfolio selection, asset pricing, and contingent claim pricing; see e.g. Friedman (2002a, 2002b), Calafiore (2007), Barillas et al. (2009), Hansen and Sargent (2008) and others cited there. Special instances of the maximum loss theorem discussed below have been used already in Friedman (2002a) and Hansen and Sargent (2008), who considered linear and quadratic portfolios depending on normally distributed risk factors.

4.2 Relative Entropy and Estimation Errors

A partial analysis of model risk often addresses parameter estimation errors but assumes the model class to be well specified. Parameter estimation errors may lead to a distribution differing from the true one in mean, correlations, volatilities, etc. Which range of distribution parameters is plausible enough to be considered in a model risk analysis? If the model class is an exponential family, the confidence regions are specified in terms of relative entropy. Many of the common distributions of statistical interest are of the exponential type: normal, χ^2, Poisson, binomial, multinomial, negative binomial, etc.

4.3 Maximum Loss Theorem

Choosing for Γ relative entropy balls of radius *k* around the distribution *v*, problem (1) reads

$$\mathrm{MR}(\mathrm{L},\mathrm{k}) := \sup_{\mathrm{Q}:\mathrm{D}(\mathrm{Q}\|v)\leq \mathrm{k}} \mathrm{E}_{\mathrm{Q}}(\mathrm{L}). \tag{3}$$

Breuer and Csiszár (2012) solved this problem explicitly by translating the problem into the problem of solving an integral equation in one variable, which under some regularity conditions has a unique solution. This solution determines the worst-case distribution in and the maximum expected loss over distributions in Γ.

This generalizes the results about the most severe scenario among a set of plausible point scenarios, which were described in section 3. Yet while this setting provides for the identification of models with risk factor distributions, it at the same time neglects errors in the specification of the loss function *L*, which are an important aspect of model risk.

Example 4. *For the linear portfolio of example 1, the worst-case scenario is a normal distribution with the same covariance matrix* Σ *as the reference distribution v, but with the mean equaling* $\mu - \frac{h}{\sqrt{l^T \Sigma l}} \Sigma l$, *where* $h=\sqrt{2k}$. *The worst-case loss is* $MR(L,k)=\sqrt{2k}\,\sqrt{l^T \Sigma l}$,

[3] *Distance is written in quotation marks because relative entropy is not strictly speaking a distance measure because it is neither symmetric nor does it fulfill the triangle inequality.*

which equals the loss in the worst pure scenario over the ellipsoid

$$\{\mathbf{r} : \sqrt{(\omega-\mu)^T \Sigma^{-1} (\omega-\mu)} \leq h\}.$$

4.4 Systematic Stress Testing and Decision Theory

Expression (1) also plays a role in decision theory. It allows us to interpret the choice of a portfolio – or measures to rebalance or hedge assets – resulting in a portfolio with a loss-pricing function L as the move of an ambiguity-averse decision-maker with multiple priors (see Gilboa and Schmeidler (1989), Casadesus-Masanell et al. (2000) – also known as maxmin expected utility (MMEU) theory). According to MMEU, ambiguity-averse agents prefer acts with lower values of (1). The set Γ is interpreted as a set of priors held by the agent, and ambiguity is reflected by the multiplicity of the priors. A decision-maker who ranks portfolios by lower values of L is ambiguity averse. And vice versa: Ambiguity-averse decision-makers act as if they were minimizing the loss function L. The relation (1) between ambiguity and risk has to be fleshed out by specifying the set Γ.

4.5 Stress Tests as Risk Measures

Expression (1), the worst expected loss over a fixed Γ of scenarios, also defines a coherent risk measure. It could thus serve as a stress test-based capital requirement. And what is more, *any* coherent risk measure can be represented as the worst expected loss over an appropriate set of risk factor distributions (see Artzner et al., 1998).

5 Stress Tests and Pricing Model Risk

Stress tests use models of the loss-pricing function L. Such models describe the loss of the portfolio as a function of the specified risk factors. Typically the number of risk factors in the model, although it may go into the thousands, is much smaller than the number of variables influencing the loss. The risk factors are (derived from) prices of basic financial instruments. Describing the price of the portfolio as a function of the prices of these basic instruments is a modeling exercise, which is prone to errors. It involves asset pricing theories of finance with highly nontrivial assumptions on no arbitrage, complete markets, equilibrium, etc. While these asset pricing theories are widely used in business as well as in the public sector, there is yet little evidence that they explain past portfolio values very well and even less evidence that they are very good in predicting the future value of a given portfolio of financial instruments (see Bossaerts, 2002). To acknowledge this fact, a good stress-testing model should be robust with respect to the specification of the loss function.

The question of the robustness of the loss function ranges from the question of what valuation or pricing model would be the right basis for the loss function, to the question of whether prices should be marked to market or based on impact-adjusted valuation such as liquidation prices. Clearly the reliance on standard asset-pricing models combined with mark-to-market valuation was one of the main reasons why risk management systems failed dramatically in the run-up to the financial crisis.

There is a recent growing literature dealing with the effects of liquidity on pricing: Important references are Holmstrom and Tirole (1997), Brunnermeier and Pedersen (2009) or Geanakoplos (2003). Most of the papers in the economics literature, including the papers referenced above, provide a qualitative theoretical understanding of how liquidity and pricing interact. These models can, however, not be directly applied to a quantitative analysis of loss-function robustness with respect to pricing risks. While a systematic analysis of loss-function robustness might be out of reach at the moment because it is not clear which sort of perturbations should be looked at, the results in Caccioli et al. (2012) suggest that some progress has been made in arriving at a quantitative understanding of the problems involved in mark-to-market accounting. Caccioli et al. (2012) find evidence that liquidation prices are approximately

$$p_f = p_0(1 - Y\sigma\sqrt{S/V}), \qquad (4)$$

where p_0 is the current price, S is the size of the position to be liquidated, V is the daily transaction volume, Y is a numerical constant of order unity, and σ is the daily volatility. The liquidation discount increases with the size of the liquidated position in comparison to the market, and with the volatility of the price. Valuating a portfolio with liquidation-discounted prices p_f instead of mark-to-market prices p_0 brings into the picture the actions of other agents via the variables V and σ. But it falls short of modeling the feedback between the act chosen by an agent and actions of other agents, which affect the risk factor distribution of the first agent.

The valuation error that could be made by relying on mark-to-market approaches for a baseline valuation is particularly strong for leveraged positions. While the mark-to-market value of a leveraged position might be high, the liquidation values might be next to zero or even negative.

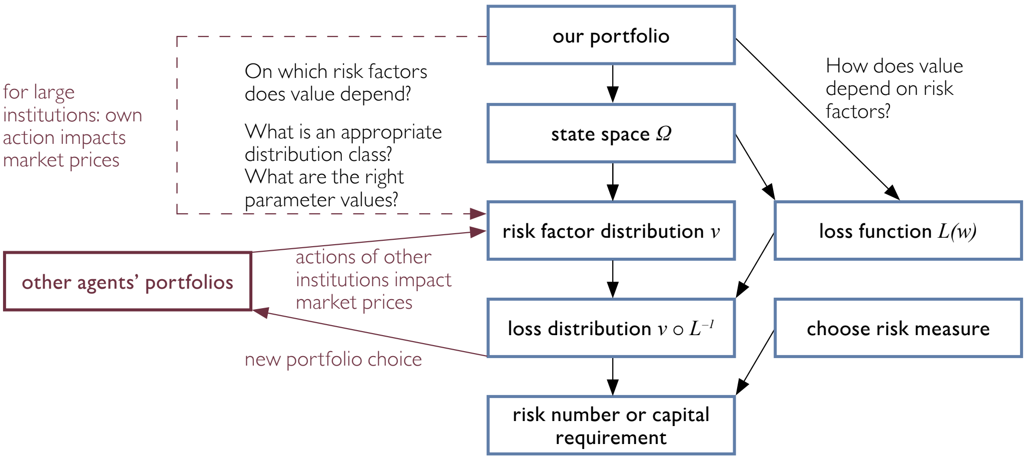
6 Open Problems

Stress testing is a relatively new field. It is therefore not surprising that it is a field with many open problems. From these many problems we would like to pick the one we believe needs particular attention from researchers: Stress situations for financial portfolios are not exclusively but typically situations of widespread distress in the financial system. The microprudential and the macroprudential perspective cannot be separated any more. Financial crisis situations and prices that emerge in such a situation are a systemic problem that can only be analyzed by understanding the interaction and feedback between individual actions and the pricing of risk in the financial system. The assumption of an exogenous risk factor distribution is inappropriate in such a situation. We must find ways to take into account the systemic nature and the endogeneity of financial risk. The models we have discussed so far ignore this risk endogeneity by thinking about risk and uncertainty in the context of a single-person decision problem, where risks are exogenous and do not depend on the behavior of individuals.

Recent accounts of financial crisis such as Brunnermeier (2009), Shin (2010) or Geanakoplos (2010) suggest the following boom-bust narrative of financial distress: A crisis usually begins in good times. People become more optimistic and get convinced that fundamental structural changes in the economy would allow taking on greater financial risk. This is usually the beginning of a leverage cycle, in which lenders lower their lending standards and collateral requirements, allowing the most upbeat investors to conduct leveraged asset purchases, feeding an asset price boom. If risks are measured from historical data, perceived risk decreases, allowing for yet more leverage in the system. At the peak of a leverage cycle, minor decreases in asset values can drive the most leveraged investors into default. The assets serving as collateral go to other investors who value the assets less highly, reinforcing the

Chart 2

Standard Risk Measurement Procedure



Source: Authors' analysis.

decline in asset prices and potentially driving more investors into default. At this stage lenders step up their lending standards and liquidity evaporates, forcing fire sales of leveraged institutions and individuals, feeding the negative spiral even further.

A stress-testing model that can take these self-feeding boom-bust scenarios into account would need to depart significantly from the current stress-testing framework that we have discussed in this paper. Rather than conceptualizing the stress test as a single-person decision problem, we need to think of the stress test as an interactive decision problem, in which certain risks result from the interaction of individuals or institutions.

To better understand how a situation of systemic risk changes the standard stress-testing framework, let us revert to the structure of the standard stress-testing model. With endogenous risk, three additional interaction channels enter the picture (chart 2). While in the standard framework the influence of my portfolio choice on the risk factor distribution is ignored, this influence has to be taken into account in an endogenous risk framework. This channel does not appear in the standard framework because it is assumed that the institution which does the stress test is negligible in the financial system as a whole and can treat risk factor distributions as given. Even when it is assumed that my own portfolio is fixed and given, in an endogenous risk world the risk factor distribution depends on all the other portfolio decisions in the system. This is also ignored in the standard stress-testing framework. Additionally it has to be assumed that, even if I keep my portfolio constant, other participants in the system will adjust their portfolios. Thus there is a feedback loop between the risk factor distribution and the portfolio composition of all the other participants in the financial system. This feedback loop is also ignored in the traditional stress-testing framework.

Some recent papers suggest different variations to the traditional stress-testing framework without developing a systematic general analysis of risk assessment and stress testing in a world with endogenous risk. The papers either depart in suggesting the consideration of a wider set of scenarios than the standard scenarios that have been used in the past. An example for these kinds of suggestions is Greenlaw et al. (2012). Others differ with regard to their suggestions of how the loss function can be reformulated to more accurately capture the systemic nature of the risks. Papers in this direction include for instance Elsinger et al. (2006a, 2006b), Gauthier et al. (2012), Duffie (2011) and Pritsker (2012) as well as Acharya et al. (2012). Finally Brunnermeier et al. (2011) suggest new approaches to data collection that would in principle allow the development of a stress-testing framework based on interactive decisions and endogenous risk.

Despite the progress made in these papers, we believe that we still lack a stress-testing framework which seriously takes into account the endogeneity of risk factor distribution. In this context there is an interesting connection of the endogenous risk problem to the paper by Caccioli et al. (2012). The empirical regularities that are summarized in a universal impact function might be a bridge to the problem of how positions should be valued in a system where endogenous risk is prevalent without modeling a fully-fledged interactive decision problem.

Cont and Wagalath (2012) propose a way to quantify the influence of fire sales on both prices and risk factor distribution. Starting from assumed deleveraging schedules for banks, and assuming that assets are sold proportionally in the deleveraging process, they show that realized correlations between returns of assets increase further in bad scenarios, due to deleveraging. Such an approach could be the basis of stress test procedures taking into account the endogeneity of risk and feedback effects of market participants' reaction to adverse scenarios.

7 Conclusions

Stress testing is a new field and as a result there is still scope for methodological improvement. Given the high uncertainty that goes with the task of making quantitative assessments about the risk-bearing capacity of financial institutions under extremely adverse circumstances, model robustness is a highly desirable property. As reviewed in this paper, the biggest advances in enhancing the robustness of stress-testing models have been made with respect to the assumptions about risk factor distribution. The key idea is not to rely exclusively on the estimated risk factor distribution but to consider a larger set of distributions as possible. In this context a coherent and universally applicable stress-testing model can be formulated that generalizes the main ideas of worst-case search over plausible domains of distributions. It helps to make stress tests robust not only with respect to distributional assumptions and gives precise meaning to the requirements that a stress test should consider scenarios that are extreme yet plausible.

We have seen that risk factor distribution is not the only model-vulnerable input to a stress test. The other object that is of concern for designing robust procedures is the loss function. To our best knowledge there has yet been no work on loss function robustness in a stress-testing environment. Given the progress made in assessing the problems of mark-to-market valuation, which could be useful for enhancing loss function robustness, this area would be an interesting avenue for deeper investigation.

Mark-to-market pricing versus liquidation-discounted pricing as in equation (4) is an issue that could build the bridge to the biggest unresolved problem in current stress testing. Although the use of a liquidation discount falls short of a full-fledged model of the feedback between the actions of different agents and the risk factor distribution, it provides a simple description of some consequences of this feedback: stronger than normal price declines when agents deleverage. Using liquidation-discounted prices instead of mark-to-market prices would allow applying the more traditional stress-testing techniques discussed here, and still take into account some of the more important consequences of risk generated by agents within the system.

References

Acharya, V., L. Pedersen, T. Philippon and M. Richardson. 2012. Measuring systemic risk. Working paper.

Ali, S. M. and S. D. Silvey. 1966. A general class of coefficients of divergence of one distribution from another. In: Journal of the Royal Statistical Society (Series B) 28. 131–142.

Artzner, P., F. Delbaen, J. M. Eber and D. Heath. 1998. Coherent measures of risk. In: Mathematical Finance 9. 203–228. www.math.ethz.ch/~delbaen/ftp/preprints/CoherentMF.pdf (retrieved on April 18, 2013).

Avellaneda, M. and A. Paras. 1996. Managing the volatility risk of portfolios of derivative securities: the Lagrangian uncertain volatility model. In: Applied Mathematical Finance 3. 21–52.

Avellaneda, M., C. Friedman, R. Holmes and D. Samperi. 1997. Calibrating volatility surfaces via relative entropy minimization. In: Applied Mathematical Finance 4. 37–64.

Barrillas, F., L. P. Hansen and T. J. Sargent. 2009. Doubts or variability? In: Journal of Economic Theory 144. 2388–2418.

Basel Committee on Banking Supervision. 2009. Principles for sound stress testing practices and supervision. BIS Technical report.
www.bis.org/publ/bcbs155.pdf (retrieved on April 18, 2013).

Borwein, J., R. Choksi and P. Maréchal. 2003. Probability distributions of assets inferred from option prices via the principle of maximum entropy. In: SIAM Journal of Optimization 12. 464–478.

Bossaerts, P. 2002. The Paradox of Asset Pricing. Princeton University Press.

Bregman, L. M. 1967. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. In: USSR Computational Mathematics and Mathematical Physics. 200–217.

Breuer, T. 2008. Overcoming dimensional dependence of worst case scenarios and maximum loss. In: Journal of Risk 110(1). 79–92.

Breuer, T., M. Jandacka, J. Mencía and M. Summer. 2012. A systematic approach to multiperiod stress testing of portfolio credit risk. In: Journal of Banking and Finance 360(2). 332–340.

Breuer, T. and I. Csiszár. 2013. Systematic stress tests with entropic plausibility constraints. In: Journal of Banking and Finance 37. 1552–1559.

Breuer, T. and G. Krenn. 1999. Stress testing. OeNB Guidelines on Market Risk 5.
www.oenb.at/en/img/band5ev40_tcm16-20475.pdf (retrieved on April 18, 2013).

Brunnermeier, M. 2009. Deciphering the liquidity and credit crunch 2007–2008. In: Journal of Economic Perspectives 230(1). 77–100.

Brunnermeier, M. and L. Pedersen. 2009. Market liquidity and funding liquidity. In: Review of Financial Studies 22. 2201–2238.

Brunnermeier, M., G. Gorton and A. Krishnamurty. 2011. Risk topography. Working paper. www.princeton.edu/~markus/research/papers/risk_topography.pdf
(retrieved on April 18, 2013).

Caccioli, F., J. P. Bouchaud and D. Farmer. 2012. Impact-adjusted valuation and the criticality of leverage. Technical report. arXiv:1204.0922v1.

Calafiore, G. C. 2007. Ambiguous risk measures and optimal robust portfolios. In: SIAM Journal of Optimization 18. 853–877.

Casadesus-Masanell, R., P. Klibanoff and E. Ozdenorem. 2000. Maxmin expected utility of Savage acts with a set of priors. In: Journal of Economic Theory 92. 35–65.

Cont, R. and L. Wagalath. 2012. Running for the exit: Distressed selling and endogenous correlation in financial markets. In: Mathematical Finance. DOI: 10.1111/j.1467-9965.2011.00510.x.

Crouhy, M., D. Galai and R. Mark. 1998. Model risk. In: Journal of Financial Engineering 7. 267–288.

Csiszár, I. 1963. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. In: Publications of the Mathematical Institute of the Hungarian Academy of Sciences 8. 85–108.

Csiszár, I. 1967. Information-type measures of difference of probability distributions and indirect observations. In: Studia Scientiarum Mathematicarum Hungarica 2. 299–318.

Duffie, D. 2011. Systemic risk exposure: A 10-by-10-by-10 approach. NBER Working Paper 17281. www.nber.org/papers/w17281 (retrieved on April 18, 2013).

Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. In: The Quarterly Journal of Economics 75. 643–669.

Elsinger, H., A. Lehar and M. Summer. 2006a. Risk assessment for banking systems. In: Managements Science 520(9). 1301–1314.

Elsinger, H., A. Lehar and M. Summer. 2006b. Using market information for banking system risk assessment. In: International Journal of Central Banking 20(1). 137–165.

Friedman, C. 2002a. Confronting model misspecification in finance: Tractable collections of scenario probability measures for robust financial optimization problems. In: International Journal of Theoretical and Applied Finance 5. 33–54.

Friedman, C. 2002b. Conditional value-at-risk in the presence of multiple probability measures. In: Journal of Risk 4. 69–92.

Gauthier, C., A. Lehar and M. Souissi. 2012. Macroprudential capital requirements and systemic risk. In: Journal of Financial Intermediation 21(4). 594–618.

Geanakoplos, J. 2003. Liquidity, Default and Crashes: Endogenous Contracts in General Equilibrium. In: Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress. Econometric Society Monographs 2. 170–205.

Geanakoplos, J. 2010. The Leverage Cycle. In: NBER Macroeconomics Annual 2009. 1–65. www.nber.org/chapters/c11786. (retrieved on April 18, 2013)

Gibson, R. 2000. Model Risk: Concepts, Calibration and Pricing. Risk Books.

Gilboa, I. and D. Schmeidler. 1989. Maxmin expected utility with non-unique prior. In: Journal of Mathematical Economics 18. 141–153.

Golan, A., G. G. Judge and D. Miller. 1996. Maximum Entropy Econometrics: Robust Estimation with Limited Data. Wiley.

Greenlaw, D., A. Kashyap, K. Schoenholtz and H. S. Shin. 2012. Stressed out: Macroprudential principles for stress testing. Chicago Booth Research Paper 12–08. ssrn.com/abstract=2004380.

Hansen, L. P. and T. J. Sargent. 2001. Robust control and model uncertainty. In: American Economic Review 91. 60–66.

Hansen, L. P. and T. J. Sargent. 2008. Robustness. Princeton University Press.

Holmstrom, B. and J. Tirole. 1997. Financial intermediation, loanable funds and the real sector. In: Quarterly Journal of Economics 112. 663–692.

Jorion, P. 2000 (2nd ed.). Value-at-Risk. McGraw-Hill.

Kerkhof, J., B. Melenberg and H. Schumacher. 2010. Model risk and capital reserves. In: Journal of Banking and Finance 34. 267–279.

Pritsker, M. 2012. Enhanced stress testing and financial stability. 2012. Federal Reserve Bank of Boston Working Paper.

Shin, H. S. 2010. Risk and Liquidity. Oxford University Press. Clarendon Lectures in Finance.

Studer, G. 1997. Maximum Loss for Measurement of Market Risk. ETH Zurich dissertation. Also available from www.gloriamundi.org/picsresources/gsmlm.pdf.

Studer, G. 1999. Market risk computation for nonlinear portfolios. In: Journal of Risk 10(4). 33–53.