IMF Working Paper

Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward

by Dimitri G. Demekas

INTERNATIONAL MONETARY FUND

IMF Working Paper

Monetary and Capital Markets Department

Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward

Prepared by Dimitri G. Demekas[1]

June 2015

Abstract

Giving stress tests a macroprudential perspective requires (i) incorporating general equilibrium dimensions, so that the outcome of the test depends not only on the size of the shock and the buffers of individual institutions but also on their behavioral responses and their interactions with each other and with other economic agents; and (ii) focusing on the resilience of the system as a whole. Progress has been made toward the first goal: several models are now available that attempt to integrate solvency, liquidity, and other sources of risk and to capture some behavioral responses and feedback effects. But building models that measure correctly systemic risk and the contribution of individual institutions to it while, at the same time, relating the results to the established regulatory framework has proved more difficult. Looking forward, making macroprudential stress tests more effective would entail using a variety of analytical approaches and scenarios, integrating non-bank financial entities, and exploring the use of agent-based models. As well, macroprudential stress tests should not be used in isolation but be treated as complements to other tools and—crucially—be combined with microprudential perspectives.

Contents   **Page**

I. WHAT ARE MACROPRUDENTIAL STRESS TESTS?

Stress testing is not a recent invention. Originally used in engineering, stress analysis is a technique of testing a structure or system beyond normal operating capacity, often to breaking point, to confirm specifications are met, determine breaking limits, or examine modes of failure. Asset managers and financial institutions, as well as their supervisors, have also realized the benefit of submitting portfolios or entire balance sheets to numerical simulations of hypothetical "shocks" to selected variables, like various asset prices, and assessing the impact on profits, capital, or the ability of regulated institutions to continue meeting their obligations, including observing regulatory requirements.

One of the early adopters of stress tests in the early 1990s was J.P. Morgan, whose RiskMetrics methodology used Value at Risk (VaR) to measure market risk (Zangari 1996):[2] in other words, the potential loss over a specified time period from movements in asset prices with a certain probability (e.g., a 5 percent one month value at risk of $100 million). Regulators caught up after a while, and the Basel II capital framework required banks to perform stressed VaR tests for market risk and, in some cases, credit risk, as part of minimum capital regulation (Pillar 1).[3] Additional stress tests could be required in the context of Pillar 2, which provided supervisors with powers to require management action by individual banks, if deemed necessary. These tests had to be *plausible*, *severe*, and *relevant* in helping the bank evaluate its capacity to absorb potentially large losses and identify steps it can take to reduce risk and conserve capital (BCBS 2005).

These stress tests had a microprudential focus: their objective was to assess the likelihood of failure of an individual institution under adverse conditions. This was consistent with the dominant approach to financial regulation at the time. It was well understood that banks financing themselves with insured deposits have an incentive to take excessive risks. Therefore, the goal of capital regulation was to force banks to internalize (some of) the unexpected losses should these risks materialize, thus mitigating moral hazard and protecting the deposit insurance fund. Stress testing was a way to assess potential losses under adverse scenarios, hence a key input to capital regulation. Equipped with this tool, regulators could ensure the soundness of each individual institution by requiring it to hold a minimum amount of capital in proportion to its risky assets. And this, in turn, would ensure the stability of the financial system as a whole.

But even as bank regulators were putting the finishing touches on the Basel II framework, many understood that this approach—especially the last proposition—was problematic: ensuring the soundness of each individual institution was neither necessary nor sufficient to ensure that the financial system as a whole would remain stable and continue to provide

---

[2] This service was so successful with its clients that J.P. Morgan spun off its Corporate Risk Management Department as RiskMetrics Group in 1998. RiskMetrics was acquired by MSCI in 2010.

[3] The use of stressed VaR, which takes into account a one-year period relating of significant losses rather than the most recent one-year period, aimed at reducing the procyclicality of capital requirements for market risk (BCBS 2009).

financial intermediation services to the economy. As Andrew Crockett, then General Manager of the Bank for International Settlements, put it: the microprudential approach to financial regulation may "strive for too much and deliver too little." It may strive for too much because the occasional failure of individual institutions is not the problem, if other institutions are capable of stepping in and providing intermediation services; trying to avoid such outcomes risks providing "excessive protection." And it may deliver too little because it does not take into account *how* each individual institution pursues compliance with capital regulation. When, for example, a regulator pushes a troubled bank to restore its capital ratio, the regulator does not care whether the bank increases its capital or shrinks assets. But if a substantial proportion of the financial system shrinks assets simultaneously to meet capital requirements, the damage to the economy may be considerable. Unless the regulators take into account the collective behavior of institutions in response to a shock (or to regulatory requirements) and its possible impact on the financial system and the economy, they may fail to minimize the probability of distress for the system as a whole and the associated economic costs—in short, systemic risk (Crockett 2000; see also Borio 2003, Kashyap and Stein 2004).

The most recent global financial crisis underscored dramatically the importance of systemic risk, as well as the failure of microprudential regulation to contain it. In his often-quoted statement in 2008, U.S. Federal Reserve Chairman Ben Bernanke called for a widening of the "field of vision" of regulators and supervisors to incorporate systemic risk (Bernanke 2008). Or, as Andrew Crockett had put it, "marrying the microprudential and macroprudential dimensions of financial stability."

What exactly does this mean for stress tests? To be sure, the aftermath of the global financial crisis saw an explosion of interest in stress tests, which have now become a prominent—and in some cases statutory—feature of regulatory regimes in many jurisdictions. Stress testing, once an arcane subject, has become a household name. But how do macroprudential stress tests differ from the traditional stress tests hitherto used for microprudential oversight? And how far has this "wider field of vision" been adopted in practice?

Different researchers have approached these questions from different perspectives.

- Drawing on the immediate lessons from the recent crisis, which highlighted the importance of liquidity and funding markets, a number of researchers pointed out that by being exclusively focused on solvency or, at best, treating solvency and liquidity separately, microprudential stress tests ignored their intrinsic interconnections (see, for example, IMF 2008; Gorton and Metrick 2009; Afonso at al. 2010; BCBS 2013). This spurred a race to develop stress testing models integrating solvency and liquidity—and eventually other types of—risk.

- Others argued that the solvency of individual banks—the focus of microprudential stress tests—was not a good proxy for systemic risk. Instead, they proposed starting from a aggregate metric of systemic risk and measuring the marginal contribution of individual banks to this metric. These approaches include the distress dependence model by Segoviano and Goodhart 2009; the risk budgeting approach (Huang et al.

2009; Chan-Lau 2010; Gauthier et al. 2010; Tarashev et al. 2010); the conditional VaR (CoVaR) by Adrian and Brunnermeier 2011; and SRISK by Acharya et al. 2013.

- Still others took a more holistic approach and put forward general principles for the design and implementation of macroprudential stress tests (Greenlaw et al. 2012; Borio et al. 2012; IMF 2012a).

- In evaluating its own post-crisis experience, the U.S. Federal Reserve highlighted the impact of common exposures, the effects of possible fire sales, the risk of reduced access to funding in stressed environments, and the importance of looking horizontally across banks rather than at a single institution at a time as the key macroprudential elements of stress tests (Bernanke 2013; Tarullo 2014).

Though far from exhaustive, this brief overview is suggestive of the essential differences between the traditional microprudential and the "new generation" macroprudential stress tests. These differences can be boiled down into two:

- Introducing *general equilibrium dimensions*, so that the outcome of the stress tests depends not only on the size and nature of the initial shock and the buffers of individual financial institutions but also on the behavioral responses of these institutions as the shock unfolds and on the interactions of these institutions with each other and with other economic agents (borrowers, funding providers, depositors).

- Shifting the focus of the stress tests from individual institutions to the *resilience of the system as a whole*, in other words on its ability to continue functioning and providing financial intermediation services to the economy.

The rest of this paper discusses *first*, how much progress has been made in addressing these two challenges in practice; and *second*, what should be the key priorities going forward in order to complete this work and establish Bernanke's "wider field of vision" in stress testing. The second, more prescriptive, part of the paper inevitably reflects the author's subjective views, informed by his experience at the IMF with stress testing tools and practices in a number of different countries.

Two important caveats are worth emphasizing upfront. First, in surveying the international experience, the paper focuses on stress testing frameworks actually used by financial supervisors, macroprudential authorities, and international organizations. It is not a survey of the academic literature on macrofinancial modeling—although advances in macrofinancial modeling are often the precursor of improvements in the models used by stress testers in the real world. Second, this paper is not a comprehensive survey of stress testing practices across the board nor does it purport to provide advice on how to run "better" stress tests overall. The focus is solely the macroprudential aspect of stress testing. Therefore, important aspects that affect the quality and effectiveness of both micro- and macroprudential stress tests, like scenario selection and design, shock calibration, transparency and communication of results, or policy follow-up, are not covered here. Fortunately, there is no shortage of studies covering these other aspects of stress tests.

II. IMPLEMENTING MACROPRUDENTIAL STRESS TESTS: PROGRESS SO FAR

A. Introducing General Equilibrium Dimensions in Stress Tests

All the reasons put forward in the literature to explain why banks, left to their own devices, will not provide the socially optimal amount of financial stability are grounded in the behavioral responses of banks to shocks and the possibility of contagion through interconnectedness. To cite just a few, these include the fear of fire sales that turns banks into "illiquidity seekers" at times of stress (Diamond and Rajan 2011); contagion through imperfect networks (Allen and Gale 2007); imperfect information about other banks' risk exposures (as, for example, in the models of Bolton et al. 2009 or Heider et al. 2009); and incentives for banks in distress to fail together so as to increase the likelihood they will be bailed out (Acharya and Yorulmazer 2008). More recently, empirical work has confirmed that bank funding costs are correlated with bank capital as a result of the interconnections between bank solvency and the expectations of market participants (Pierret 2014; Aymanns et al. 2015).

Against this background, the partial equilibrium approach of traditional microprudential stress testing models looks relatively primitive. It consists in assuming an exogenous shock, for instance on credit quality or asset prices, and tracing the impact of associated losses on the capital position of each individual bank, making simplistic assumptions about its behavior following the shock (on profit distribution, credit expansion or deleveraging, etc.). Liquidity and market risk are treated separately from solvency, if at all, and interactions among banks and macro feedback effects are generally ignored.

The shortcomings of this partial equilibrium approach are more glaring the longer the time horizon covered by the tests. For example, a static balance sheet may be an acceptable simplifying assumption in stress tests covering one year, but becomes a severe handicap in tests covering a period of three or five years after the initial shock. And with stress tests being increasingly based on macroeconomic scenarios, in which the lags between the original shock and the full effect on credit quality and the bank balance sheet can be considerable, their time horizon has been lengthening.

Not surprisingly, these shortcomings came under the stress testers' microscope early on, and substantial intellectual energy has been expended in trying to address them, so as to move away from the partial equilibrium approach of microprudential stress tests and introduce more system-wide or general equilibrium dimensions.

Before reviewing the progress in this area, it may be worth stopping to ask: how much general equilibrium is too much? The question may seem surprising, but in fact, a "full" general equilibrium approach to stress tests, even if it were feasible, would not be desirable. To start with the most obvious point, incorporating in the stress scenario the likely policy reaction of regulators or the government, ranging from forbearance to corrective action, intervention, or bail out, would make the exercise less informative for supervisors and less likely to foster better risk management in banks. But even incorporating the full reaction of the banks themselves to the stress scenario could be problematic. Bank management actions in response to an adverse shock can range from "business as usual" to adjusting prices or

terms for certain products, increasing provisions, reducing exposures, constraining dividends, or raising additional capital. Leaving aside the difficulty of predicting and modeling these actions—and their potential general equilibrium interactions—it is clear that allowing the full range in the tests may mask the true extent of distress facing banks as a result of the assumed shock. Stress tests will thus always be partial equilibrium exercises, and there is no easy answer to the question how far to go in incorporating general equilibrium dimensions: it is a matter of judgment. Just how thorny this judgment is can be glimpsed by reading the detailed instructions of the U.S. Federal Reserve to banks for the Comprehensive Capital Analysis and Review (CCAR) exercise or the Bank of England's careful deliberations on the new framework for bank stress tests (Board of Governors of the Federal Reserve System 2014; Bank of England 2013).

Turning now to the efforts the stress testing community has made in recent years to incorporate general equilibrium dimensions in stress tests, they can be grouped in two sets of approaches.

- *Balance sheet-based models* that use balance sheet data to model the interactions between different types of risk, introduce more sophisticated treatment of bank behavioral responses to shocks, or capture the impact of interconnectedness.

- *Market price-based approaches* that use (mostly) market data and statistical techniques to capture interlinkages between institutions, markets, or sources of risk.

Like every taxonomy, this has caveats: the dividing line is drawn somewhat arbitrarily; it does not provide a neat classification for a few models that combine elements of both approaches; and in the case of the second group, in particular, it bundles together models that are conceptually quite different. Alternative taxonomies are also possible: for example, distinguishing between models that start with individual institution data (balance sheet, market price-based, or a combination of both) and build aggregate systemic risk metrics from those, and models that start at the system-wide level and then derive institution-specific measures of distress. The merits of this particular taxonomy are that models that rely (mainly) on balance sheet data or on market data share many common strengths and pitfalls; that this distinction is one that most stress test practitioners would immediately grasp; and, of course, that some taxonomy is better than none.

The next section presents some representative examples of each type of approach and discusses their relative merits and limitations. The presentation focuses mainly on frameworks actually used for stress testing by central banks and other national or international authorities and is, by necessity, selective.

Balance sheet-based models

This is an area where significant progress has been made in recent years. There is now a wealth of different stress testing models in use by central banks and supervisory agencies that go a considerable distance toward capturing at least some of the general equilibrium effects likely to have a major impact in a stress scenario.

The point of departure for these models is the basic framework used for microprudential stress tests. In this framework, individual bank balance sheet data, in different degrees of granularity, are used to assess the impact of an exogenous shock on asset quality, income, and ultimately capital (for solvency tests) or various measures of liquidity (for liquidity tests) of each individual bank. A number of contributions have expanded this familiar framework to integrate these two, as well as other, types of risk and relax the simplistic assumptions about bank behavioral responses to the shock.

- o Barnhill and Schumacher (2011) model jointly systemic liquidity and solvency risks in a simple framework where systemic liquidity shocks ("runs") are modeled as a reaction to shocks to asset values that drive bank capital ratios below a certain critical level. In this framework, runs are essentially an extreme form of market discipline on banks by those providing funding (depositors, wholesale investors, and other banks).

- o The Bank of England was one of the pioneers in developing a model that integrates credit, market, and liquidity risk (the Risk Assessment Model of Systemic Institutions —RAMSI; see Alessandri et al. 2009; Aikman et al. 2009). In this model, a macro shock leads to both credit and market losses, which then lead to a downgrade in the bank's liquidity through a simple scoring system. In turn, this causes higher funding costs. RAMSI also departs from the standard static balance sheet assumption by allowing banks to change asset composition in response to losses, thus incorporating to some extent banks' behavioral responses in the shock scenario.

- o Building on the RAMSI, Kapadia et al. (2012) propose a model in which funding distress for individual banks can arise not only as a result of solvency concerns but as a function of a broader set of indicators that proxy solvency, liquidity profile, and confidence. This allows for liquidity shocks to arise even if the bank remains solvent.

- o Another pioneer in developing integrated frameworks for solvency and liquidity risk was the Oesterreichische Nationalbank with its Applied Risk, Network, and Impact Assessment Engine (ARNIE) (Feldkircher et al. 2013; Puhr and Schmitz 2013). ARNIE includes a detailed model of bank cash flows, distinguishing between contractual and behavioral flows, and allows for two-way interactions between solvency and liquidity via multiple channels: (i) the rating migrations of asset classes in the solvency tests affect the quality of collateral available for deposit at the central bank; (ii) defaulting assets reduce cash inflows; (iii) the solvency position of the bank affects its access to funding markets; and (iv) hikes in the cost of funding or losses from asset fire sales influence solvency through the profit-and-loss account. ARNIE also extends the dynamic treatment of bank balance sheets during the shock period under certain assumptions.

- o A model developed at the European Central Bank (ECB) goes even further in introducing dynamic balance sheet effects by using an optimization framework in which banks adjust their asset structure in response to changes in the expected loss of their asset categories (Henry et al. 2013). The optimization process is set up as a mean-variance portfolio choice with risky funding sources and regulatory limits that

require a bank to keep the capital ratio above a certain level and to hold a certain amount of liquid assets. Given the complexity of the model, the balance sheet is highly stylized and the model covers just one period. Covering a longer period would require either a dynamic optimization setup allowing for inter-temporal asset allocation decisions; or sequential runs of the model, with the end-period asset structure in period t set as the initial structure in period $t + 1$.

o The Bank of Canada's Macrofinancial Risk Assessment Framework (MFRAF) includes three modules that account for three different risks: credit losses as a result of a macro shock; liquidity risk if investors develop concerns about a bank's solvency and/or its liquidity position; and network spillover effects if some other banks fail to repay their interbank counterparties. MFRAF is one of several approaches that use network models as an "overlay" to the balance sheet-based stress testing framework to trace contagion through interbank exposures. A striking finding of this model is that when liquidity and network spillover effects are considered together with credit losses for a given shock, the aggregate capital position of banks declines by an additional 20 percent (Anand et al. 2014). This underscores the importance of considering these risks jointly.

o The Bank of Korea and the Bank of Japan have developed frameworks that not only integrate solvency and funding risks but also include some macro feedback effects. The Systemic Risk Assessment Model for Macroprudential Policy (SAMP) developed by the Bank of Korea has six modules tracing the impact of a macro shock on solvency and liquidity through fire sales and liquidity withdrawals, the latter using a network model of interbank exposures. In this model, a credit crunch as a result of bank solvency or liquidity pressures leads to further increases in credit default rates that feed back on bank losses (Bank of Korea 2012). The Bank of Japan's approach is similar: it uses a Financial Macroeconometric Model (FMM) to model the interactions between the real and financial sectors and use it for stress tests with a time horizon of three years (Kitamura et al. 2014).

o A model developed by researchers at the Hong Kong Monetary Authority (Wong and Hui 2009) follows a similar approach in integrating solvency and liquidity risk, but adds a twist: the default risk of an individual bank depends on the market value of its total assets, calculated through a Merton-type model. Merton-type models, discussed more extensively in the next section, build on the Black-Scholes-Merton view of the equity of a firm as a call option on its assets with the strike price being equal to its liabilities, and use market data to calculate probabilities of default, defined as the probability of the total value of liabilities exceeding the total value of assets. This captures market perceptions of the default risk of individual institutions regardless of the particular risk source. This is one of the few models that combine elements of balance sheet-based and market price-based approaches.

o Instead of explicitly modeling the various channels of interaction between solvency and funding risk, an alternative is to introduce variable hurdle rates in standard solvency tests that depend on the post-shock capital position of the bank (the hurdle rate is the level of capital a bank has to meet or exceed at the end of the stress test

period in order to "pass" the test). The intuition here is that if a bank's capital position deteriorates, providers of unsecured funding will demand higher interest rates, thus squeezing the bank's net interest margin. To avoid this outcome and ensure uninterrupted access to funding, banks should therefore target a capital ratio higher than the regulatory minimum—the hurdle rate traditionally used in microprudential stress tests. This intuition is not new: it was already pointed out in IMF (2012a) and Greenlaw et al. (2012), and was one of the main reasons for setting hurdle rates above the regulatory minima in the recent stress tests in the European Union (EBA 2014). The problem until now was that there was little hard empirical evidence to back it up, making the choice of hurdle rates other than regulatory minima somewhat arbitrary. Aymanns et al. (2015) fill this gap by examining a very large sample of banks (approximately 10,000 U.S. banks over 21 years and 2,700 global banks over 10 years) and estimating a significant negative relationship between bank solvency and (wholesale and total) funding cost. Furthermore, this relationship is nonlinear (the sensitivity of funding cost to solvency is greater at low levels of capital) and procyclical (the sensitivity is stronger during bad times).

The major benefit of these models is tractability. They all share, to a greater or lesser extent, a key feature: they rely on detailed balance sheet data of individual institutions and model explicitly the general equilibrium dimensions they intend to capture, be they solvency-liquidity interactions, behavioral responses, or macro feedback effects. This way, once the model is run, it is in principle possible to trace the impact of the shock through the various channels. This makes it feasible to attribute the overall impact of the shock to its constituent parts, measure the contribution of cross-institution contagion, and even estimate the size of second-round feedback effects if these are built into the model. It also makes it possible to assess whether individual bank responses to shocks contribute to financial stability or are collectively detrimental to it.

This benefit, however, comes at a price.

- First, these models can only capture general equilibrium effects that are explicitly built into them. Risk co-dependence, contagion, or feedback effects that are not explicitly modeled are ignored or captured only indirectly. This feature also helps explain the multiplicity of models that fall in this category, as stress testers make different choices about what system-wide effects to target with their models.

- Second, analytical and computational complexity and data requirements increase very rapidly as more features are added to the models. Moreover, given the heterogeneous dimensionality of the underlying data for the different elements covered by these models (e.g., bank credit risk, bank liability structure, interbank exposures, corporate or household balance sheets, etc.), these frameworks are typically implemented in modules. This renders them slow, cumbersome, and costly to construct and run; constrains the analysis of endogenous dynamics of shock propagation; undermines the internal coherence of stress scenarios; and sometimes requires difficult trade-offs between efficiency and granularity. Being slow and cumbersome also means that they are less effective as high-frequency monitoring tools for systemic risk: if weeks or

months have to elapse between the data collection cutoff date and the final stress test results, it is unlikely that this tool can be used more than once or twice a year.

- Third, since they rely on bank balance sheet data, they depend crucially on the availability and quality of these data. Both are problematic. For example, over-the-counter derivatives and the activities of shadow banking entities leave gaps that imply that certain bank exposures are not captured properly in the available data. As regards data quality, this is influenced by a host of factors, including internal bank controls, data reporting protocols and, crucially, the regulatory environment and supervisory effectiveness. The importance of this point was underscored by the EU decision to conduct a comprehensive balance sheet assessment before the first stress tests under the single supervisory mechanism. Reconciling accounting and economic definitions of losses, provisions, and default may also present challenges.

- Fourth, to the extent these balance sheet data come from supervisory sources and are not publicly available, the stress tests cannot be replicated by individual banks or private sector analysts. This weakens somewhat the claim that stress tests promote transparency and a more open conversation about financial stability.

A final point worth making is that, with few exceptions, the microfoundations of these models are weak. They incorporate elements that account for dynamic effects, behavioral responses, and risk integration by imposing ad hoc mechanisms and behavioral rules-of-thumb rather than an integrated framework of optimizing economic agents. Even in the model developed at the ECB—probably the most advanced from this perspective—liability optimization is absent and the loan book is treated separately from the trading portfolio. Many would argue that this is a price worth paying for introducing, in a practical way, general equilibrium dimensions in the traditional balance sheet-based stress testing models. Moreover, these behavioral rules-of-thumb are not arbitrary but derived from observations of past behavior. Nevertheless, it is an important handicap and it detracts from the usefulness of these models, especially in times of crisis, when observed behavioral patterns break down and economic agents learn and adapt, often on the basis of imperfect information.

Market price-based models

At their most basic level, these models use market information (exclusively or in combination with balance sheet data) to estimate the probability of default or distress of a bank.[4] Instead of looking at just a single indicator of bank soundness—the regulatory capital ratio—these models look at a measure that, at least in principle, reflects *all* sources of vulnerability. The relationship between this probability of default or distress and macroeconomic or other variables can, in effect, provide a stress test for the bank.

---

[4] There is, of course, a plethora of models using a combination of market and accounting data (financial ratios) to calculate the probability of default of individual firms, going back to credit scoring models or Altman's Z-score (Altman 1968). But these are not specifically designed for financial institutions or used for stress testing. And today, these simple reduced form empirical approaches have largely been displaced in the analysis of financial distress for both financial and non-financial firms by more sophisticated "structural" models, such as the Merton-type models.

Some of the most important recent academic contributions to the measurement of systemic risk—already mentioned in the first section of the paper—fall in this category. These include the risk budgeting approach, the CoVaR by Adrian and Brunnermeier 2011, and SRISK by Acharya et al. 2013. But this section focuses mainly on two specific approaches that are well established and regularly used for stress testing of banks. These are (i) models that extract from the data the multivariate density function for the banking system as a whole, modeled as a portfolio of banks, and from that calculate measures of soundness for individual banks; and (ii) models that estimate the probability of default for individual banks building on the option price analogy of the capital structure of a firm—the Merton-type models referred to earlier.

As a reminder of the caveats of taxonomy, it should be emphasized that, while sharing some common elements, these two types of models are conceptually quite different. They therefore warrant a somewhat more in-depth discussion.

Models of distress dependence

The departure point for these models is the insight that financial instability can arise as a result of systemic shocks, contagion after idiosyncratic shocks, or a combination of both (Goodhart et al. 2006). Bank distress dependence is thus a central feature of systemic risk. Moreover, contagion following a shock can arise even in the absence of direct interbank exposures (for example, through common bank exposures to third parties or investor sentiment), and the risk of contagion changes over time (it is higher at times of systemic stress). Therefore, using a "snapshot" of the network of interbank claims at a point in time to proxy bank distress dependence—a common approach in the balance sheet-based models discussed above—is far too limiting.

Segoviano and Goodhart (2009, 2010) conceptualize the financial system as a portfolio of banks (their model has recently been extended to cover non-banks—see Segoviano et al. 2015); estimate empirical measurements of individual probabilities of distress or default for each bank in the portfolio; use a non-parametric method (the Consistent Information Multivariate Density Optimizing (CIMDO) methodology[5]) to extract from these probabilities the multivariate density function for the financial system (portfolio) as a whole; and then calculate various stability metrics for individual banks. These metrics include the joint probability of distress or JPoD (the probability of *all* the banks in the system becoming distressed); the probability of *a specific bank* or group of banks becoming distressed; and the banking stability index (the expected number of banks becoming distressed conditional on one bank in the system becoming distressed). Because they are derived from the multivariate density function, these metrics embody both linear and non-linear distress dependence among

---

[5] The CIMDO methodology is based on the minimum cross-entropy approach, where a posterior multivariate distribution—the CIMDO copula function—is recovered using an optimization procedure by which a prior density function is updated with empirical information via a set of constraints. In this implementation, the empirical estimates of the probability of distress of individual banks act as the constraints, and the derived CIMDO copula is the posterior density that is the *closest* to the prior distribution and *consistent* with these constraints. This methodology and its advantages relative to other parametric multivariate densities are presented in detail in Segoviano (2006).

banks and are time-varying. They are thus superior to estimates from traditional risk models that, generally speaking, capture only linear dependence (correlations) and are assumed to remain constant for a fixed period of time.

It should be emphasized that the individual probabilities of distress or default used as input to the calculation of the multivariate density function are exogenous to the model. Any methodology could in principle be used to estimate them, including balance sheet-based methodologies. Using market data is thus not a necessary feature of this model. In practice, however, as the time-varying nature of bank distress dependence is a key strength of this approach, high-frequency market data, such as CDS spreads or stock returns, are typically used to calculate the probabilities of distress or default for individual financial institutions.

This approach has been used a number of times by the IMF, in conjunction with more traditional approaches, for macroprudential stress testing in the context of FSAPs, most recently in Switzerland (IMF 2014a) and the U.S.A. (IMF 2010).

Merton-type models

The fundamental insight behind these models rests on the work of Fischer Black, Myron Scholes, and Robert Merton, for which the latter two were awarded the 1997 Nobel Prize in economics (Black and Scholes 1973; Merton 1974). The equity of the firm can be viewed as a call option held by shareholders on its total assets, where the strike price is equal to the outstanding debt owed to bondholders at maturity. Bondholders receive a put option premium in the form of a credit spread above the risk-free rate in return for holding risky debt. A market-implied default probability—the probability that the (stochastic) value of the firm's assets falls below that of its liabilities—can thus be estimated by the duration of the bond, the degree of leverage, and asset price volatility, which is, in turn, linked to (observed) share price volatility. This implied default probability is risk-neutral, since asset pricing models weigh in investors' risk aversion, but a number of different methods can be used to transform these risk-neutral probabilities into "actual" or "real-world" default probabilities. This conceptual framework can be extended to accommodate multiple debt instruments of different maturities, different types of assets, etc.[6]

This approach is used in a number of commercially available risk management tools that estimate corporate default risk, such as the Expected Default Frequency™ and RiskCalc™ models by Moody's KMV (Crosbie and Bohn 2003; Dwyer et al. 2004).[7] These models are not specifically designed for financial institutions, although they can be used for those, too.

---

[6] For textbook treatments, see Altman and Hotchkiss 2006, Gray and Malone 2008, and Chan-Lau 2013.

[7] These tools originated with the work of Stephen Kealhofer and Oldřich Vašiček who, together with John McQuown, formed KMV company in 1989. KMV was acquired by Moody's in 2002 and continued to develop these tools. Moody's KMV was renamed Moody's Analytics in 2007.

A related model, better suited for risk analysis of financial institutions, is the Contingent Claims Analysis (CCA) (Gray and Malone 2008; Gray 2010). The CCA uses the Black-Scholes-Merton framework to construct a "risk-adjusted" balance sheet of a bank. In conventional balance sheet analysis, a change in the value of assets results in a one-for-one change in book equity. Accordingly, in traditional microprudential stress tests, expected credit losses from a shock, calculated as the probability of default (PD) of the particular claim—loan or loan category, depending on the degree of granularity—times a loss-given-default (LGD) times the exposure-at-default (EAD), are aggregated across exposures and counted against bank capital. In a "risk-adjusted" balance sheet context, however, a decline in the value of assets also increases expected losses to creditors. The change in the market value of equity thus depends not only on the expected loss on the asset side but also on the degree of leverage, the volatility of assets, and investors' risk appetite.[8] Conversely, this approach can also be used to estimate the impact on the bank's borrowing costs of higher (or lower) levels of capital, changes in risk appetite, and differences in seniority between different types of liabilities (e.g., guarantees on deposits).

The CCA model can be used for macroprudential stress testing by estimating the correlation between the expected losses to creditors (or other CCA risk indicators) and various macro variables, and tracing the impact of hypothetical shocks to these variables through the "risk-adjusted" balance sheet on the default probability of the bank or the market value of equity. The model has been used often for stress testing banks in IMF FSAPs, notably in Hong Kong SAR (IMF 2014b), Israel (IMF 2012b), Sweden (IMF 2011a), the UK (IMF 2011b), and the U.S.A. (IMF 2010).

Discussion

The major strength of both types of market price-based models is that they take a broad, and arguably more realistic, view of the solvency of a bank by looking at a metric—the probability of default or distress—that, in principle, captures all sources of vulnerability and contagion. And since banks are vulnerable to self-fulfilling runs that need not be triggered by weak fundamentals, using data that incorporate market participants' perceptions to estimate the probability or default or distress is an advantage.

Another major advantage of market price-based models, especially when compared to the balance sheet-based models discussed earlier, is their computational simplicity and relatively light data requirements. By using sophisticated techniques to extract as much information as possible about the probability of default or distress from available data, market price-based models are small and easy to run. Combined with the availability of very high frequency market data, this makes these models ideal for high frequency monitoring of bank

---

[8] The conventional accounting balance sheet can be seen as a special case of the "risk-adjusted" balance sheet when uncertainty is zero, i.e., when bank assets have no volatility.

soundness.[9] The public availability of market data also means that market price-based models can—and indeed are being—used for stress testing by private sector analysts.

It is important to underscore that, at least for the two market price-based models discussed in detail in this section, computational simplicity does not come at the expense of analytical rigor. The distress dependence model uses well-established (if little known among economists) statistical techniques to extract the multivariate density function and, from that, calculate financial stability metrics; and the CCA is squarely founded in standard finance theory.

A third advantage of market price-based models is that it is straightforward to integrate non-bank institutions in the stress tests alongside banks. There are balance sheet-based models for stress testing insurance companies, but they cannot be integrated with those for banks: stress testing banks and insurance companies with the same scenarios using balance sheet-based models can therefore only be done in separate modules. The two market price-based models discussed above, on the other hand, can easily integrate these and other types of institutions simultaneously in the same framework.

At the same time, these models have pitfalls.

- The most obvious is related to their reliance on market data. *First*, these data may not be available for all institutions in all countries (e.g., in the case of not publicly listed banks), and even where they are, shallow or illiquid markets may render them uninformative. While there are ways around this problem, it is indeed a severe handicap, especially for the CCA. *Second*, market data may not capture structural breaks owing to changes in business practices or the regulatory framework. *Third*, these models use observable market measures (equity returns, CDS spreads, etc.) to infer the underlying volatility of assets, which is unobservable. But market prices are available only for items that are traded, and many bank assets and liabilities are not. *Lastly*, market data are noisy and may overestimate or underestimate risks, often without any relation to bank fundamentals. Bank default risk indicators estimated from these data may thus be excessively volatile, and may not provide a sound basis for bank management or supervisory action.[10]

- Another important pitfall is that by extracting information from market data and constructing a summary metric of bank soundness, market price-based models do not allow the stress tester to differentiate between the various factors that contribute to the bottom line (initial shock, risk inter-dependence, common exposures, cross-

---

[9] In a recent implementation of Segoviano's distress dependence model at the Bank of Israel, it was possible to update the model and calculate the full range of stability metrics for individual banks on a daily basis.

[10] This argument merits closer scrutiny. Market data often reflect information not yet known to (or fully understood by) supervisors. In addition, regardless of their relationship—or absence thereof—with bank fundamentals as these are understood by supervisors, market trends are often self-fulfilling. Therefore, dismissing these signals as "noise" may not be a smart move.

institution contagion, or market sentiment): all these factors are combined into the implied probability of default or distress generated by the model. This had led critics to dismiss these models as "black boxes." Advocates of these models, on the other hand, argue that this is a price worth paying for a comprehensive measure of bank soundness; and point out that additional tools, including balance sheet analysis, may be used to help disentangle the contributions of various risk factors on the bottom line.

- Finally, a criticism that is often leveled against market price-based models is that they are backward-looking: they reflect information embodied in market prices at a certain point in time, and cannot predict how these will change from one day to the next, much less in a stress scenario. While true, this criticism applies just as much to models that are based on a bank balance sheet snapshot from the past. Indeed, given the data reporting/collection lags and their cost and complexity, the results of stress tests using balance sheet-based models are much more likely to be outdated by the time they are available than those of market price-based models.

B. Focusing Stress Tests on the Resilience of the Financial System as a Whole

Ensuring the soundness of individual banks is neither necessary nor sufficient for ensuring financial stability. Macroprudential stress tests should therefore focus not on the soundness of individual institutions—however this is measured—but on the resilience of the system as a whole; in other words, on its ability to continue providing uninterrupted financial intermediation services to the economy in the face of shocks.

Although this principle has been well understood for a while, most stress tests conducted by major central banks, supervisory agencies, and the IMF—no matter how sophisticated the model they use—still do not live up to it. They continue to focus on individual banks and provide a binary (pass-fail) outcome for each bank under the stress scenario. To give an idea of the system-wide effect of the shock, individual results are then added and expressed in terms of (i) the number of banks and their collective share in total system assets that fail the meet the arbitrary hurdle rate (regulatory capital ratio or other), and (ii) the total amount of the shortfall.

While focusing on individual bank results is certainly useful from a supervisory point of view (and interesting to bank shareholders, the Treasury, or the deposit guarantee agency, who may be called upon to deal with a failing bank), this approach to systemic risk has two problems: the *aggregation problem* and the *robustness problem*.

- The *aggregation problem* reflects the fact that the sum of individual banks' shortfalls vis-à-vis the hurdle rate in the event of a shock is not a good proxy of systemic vulnerability. First, expressing the outcome of the stress test in a binary (pass-fail) way may make sense when the question is whether individual banks will continue to meet regulatory minima after a shock, but is not a good way to measure each bank's contribution to systemic risk, which is a continuous concept. Second—and more important—given the different ways in which banks are interconnected, the individual shortfalls in the event of a shock are not additive. Correctly aggregating

individual probabilities of default and the associated shortfalls requires some knowledge of the dependence structure between individual bank balance sheets. And as discussed earlier, this dependence structure is non-linear and tends to vary with the degree of systemic stress, i.e., to increase at times of stress.

- The *robustness problem* expresses the notion that a single stress scenario, however severe, does not provide enough information about the resilience of the system to other shocks with the same probability. A theoretically more correct approach would be to look simultaneously at all risk factors affecting the system; estimate a multidimensional region with a given probability mass—say 95 or 99 percent; and calculate the maximum loss of the system for all scenarios falling in this region.[11] This would measure the resilience of the system to *all* plausible scenarios with a probability of at least 95 or 99 percent. But this approach is hard to implement and has not so far been used in stress tests conducted by a major central bank.

Both problems have long been recognized but are tough to crack. Some of the market price-based models reviewed in the previous section do not face the aggregation problem: models like Adrian and Brunnermeier's CoVar or Segoviano and Goodhart's distress dependence start by estimating systemic risk and then derive the individual banks' contribution to it. The issue, however, with these approaches is that they do not explain how these measures translate into each individual bank's capital or liquidity requirements. This is crucial because these are ultimately the main tools micro- or macropudential regulators can use to mitigate risk, both the idiosyncratic and the systemic varieties. A model without this element, no matter how sophisticated, is unlikely to be of much practical value to policy-makers. Although there is no single, generally accepted framework that addresses these problems, a number of very promising recent contributions have tried to address one or both.

- o The distress dependence model has been extended to allow the mapping of its results into individual banks' regulatory ratios (Segoviano et al. 2015). The CIMDO multivariate density function extracted from the data is used to generate expected system-wide losses with Monte Carlo simulations. Just as the multivariate density function incorporates the time-varying dependence structure of individual entities' distress probabilities, the simulated systemic loss distribution includes losses generated not only directly but also as a result of contagion and interconnectedness. This framework can be used for stress testing in the following way: first, by correlating the probabilities of default (PDs) for individual banks to a set of macro variables; then, projecting "stressed" values for these individual bank PDs under a

---

[11] Estimating the maximum loss for a multidimensional region of a given probability mass has an undesirable property known as the dimensional dependence of maximum loss. As an example, start with a bond portfolio with risk factors consisting of two yield curves in ten currencies. One risk manager models the yield curve using seven maturity buckets and another using fifteen maturity buckets. Both choose a plausibility region of 95 percent. Breuer (2008) has shown that the second risk manager will calculate a maximum loss that is 1.4 times higher than that calculated by the first risk manager, although both look at the same portfolio and the same plausibility level. Breuer et al. (2009) have suggested a slightly different statistical concept of plausibility that is not subject to this technical pitfall.

stress scenario; using these "stressed" PDs to extract a new multivariate density function; and finally, generating a systemic loss distribution under the stress scenario. This distribution can be used to generate a systemic expected shortfall and, through the estimated Shapley Value,[12] calculate the Marginal Contribution to Systemic Risk (MCSR) of each individual institution. The MCSR represents the proportion of systemic losses that can be attributed to an individual entity taking into account its size, its risk profile, and its interconnectedness with the rest of the system. By comparing the MCSR of each individual entity to its capital buffer, it is possible to assess whether the capital is not only enough to cover the individual institution's losses but also the institution's contribution to systemic risk under the stress scenario.

o   The CCA framework has been extended to generate a measure of systemic risk that takes into account distress dependence, thereby dealing with the aggregation problem. This "systemic CCA" (Gray and Jobst 2011; Jobst and Gray 2013) starts with the risk-adjusted balance sheets for individual institutions generated under the CCA approach and estimates expected losses (and the associated changes in capital) for each.[13] These individual expected loss distributions are assumed to be "fat-tailed" and follow a Generalized Extreme Value (GEV) distribution. They are then combined with a non-parametric multivariate dependence function to generate a multivariate extreme value distribution (M-GEV) of joint expected losses of all institutions in the system.[14] The M-GEV distribution is used to derive the systemic expected shortfall, which is then used to calculate each individual institution's marginal contribution to the systemic shortfall. As with the simple CCA, stress testing using the systemic CCA involves estimating the correlation between the risk-adjusted balance sheet parameters of individual banks to various macro variables; projecting "stressed" risk-adjusted balance sheets; using these to generate "stressed" GEV distributions, M-GEV, ES, and marginal contributions to systemic loss; and comparing the latter to the starting capital of each individual institution.

o   Instead of estimating system-wide losses under a stress scenario, distributing those losses among individual banks, and then comparing the outcome to starting bank capital, Webber and Willison (2011) re-cast the problem entirely from the policy-

---

[12] The Shapley Value is a formula in cooperative game theory that assigns a unique distribution among the players of the total surplus (or loss) generated by the coalition of all players. As a solution concept, the Shapley Value has the properties of efficiency, symmetry, and linearity (for more details, see Neyman 2002).

[13] Instead of the standard CCA formula, more complex option pricing formulae may also be used—without departing from the basic Merton framework—to take into account excess skewness (a shift of the average away from the median) or kurtosis (a narrower peak and fatter tails) in the distribution of expected losses (Jobst and Gray 2013). This corresponds to the empirical observation that at times of stress, default risk is highly skewed and extreme values are more likely.

[14] In contrast to the CIMDO methodology, where the multivariate copula function is inferred from individual distributions using a cross-entropy approach without imposing a specific functional form on the dependence structure, the systemic CCA combines a specific (non-parametric) dependence function with the individual distributions to derive the multivariate copula.

maker's perspective. The policy-maker is interested in setting individual bank capital at a level that ensures systemic solvency over a given time horizon at a certain level of probability. Given the trade-off between stability and efficiency, this is set up as a constrained optimization problem, where bank capital requirements are minimized subject to a specified probabilistic systemic stability target. Nested inside the policy-maker's optimization problem is a highly simplified Merton-type model of individual banks. This model[15] could be adapted for use in stress testing the way the CCA model is, and can, in principle, both deal with the aggregation problem and provide a proper measurement of systemic resilience to (solvency) risk.

o   Another promising theoretical approach that tackles both problems is outlined in Pritsker (2014). The key innovation is a definition of systemic stability that is directly related to the level of regulatory capital of each bank: the "system assets in distress" (SAD) measure is defined as the sum of each bank's "intermediation capacity," in turn defined as a function of its total assets and capital times a vector of risk factors. A continuous measure of systemic risk is then given as the probability that SAD exceeds a pre-specified level θ ($P(\text{SAD} > \theta)$) for a given time horizon. The model then sets up a constrained stress maximization (CSM) problem to estimate the amount of capital needed for each bank so as to satisfy the constraint that $P(\text{SAD} > \theta) \leq \alpha$, where α is the regulator's systemic risk target. The CSM uses Monte Carlo simulations to estimate non-parametrically the probability density function of SAD so as to take into account distress dependence *and* cover all possible realizations of the risk factors at a certain probability level. Although highly stylized and still under development, this approach has potential for macroprudential stress testing.

This short overview will have hopefully made clear why focusing macroprudential stress tests on the resilience of the system as a whole has proved challenging. It is hard to build a model that measures correctly systemic risk and the contribution of individual institutions to it (capturing distress dependence, overcoming the aggregation problem, providing a proper definition of systemic resilience) and, at the same time, relates the results to the established regulatory framework for each individual bank. And it is even harder to make this model robust enough to use in a variety of environments and for a variety of financial institutions, as well as simple enough to explain to supervisors, bank managers, and market participants. It is therefore not surprising that actual stress testing practice has advanced relatively little in this area.

---

[15] In a similar contribution, Gauthier et al. (2010) use CoVaR to allocate capital across banks according to their contribution to systemic risk (similar to decomposing portfolio risk to each component of the portfolio). That model, however, is concerned with allocating a fixed amount of capital, not determining the optimal amount of capital for the system as a whole, as well as its distribution.

III. THE WAY FORWARD FOR MACROPRUDENTIAL STRESS TESTS

Of the two challenges involved in transforming the traditional microprudential stress testing frameworks into "true" macroprudential risk assessment tools, the stress testing community has made significant progress in tackling one but much less in dealing with the other. A large number of (balance sheet and market price-based) models that go a long way toward incorporating general equilibrium dimensions into stress tests are now available and regularly used by central banks, macroprudential authorities, and others. In contrast, very few stress testing models focus on—and measure correctly—the resilience of the financial system as a whole and its ability to continue providing financial intermediation services under stress in a way that makes the results readily actionable for individual banks and their supervisors. Crockett's call for marrying the micro- and macroprudential dimensions of financial stability is still some way away as far as stress testing practice is concerned.

Complexity is a major obstacle. The traditional approach of starting with well-established frameworks—particularly the standard balance sheet-based models used for microprudential stress testing—and adding features to them, like risk integration, behavioral responses, or feedback loops, increases quickly the analytical and computational complexity, rendering the resulting frameworks very cumbersome. And moving from that to a proper measurement of systemic risk is fraught with difficulty. Market price-based models have their own pitfalls.

The remainder of this section discusses some ways to move from where we are today toward more effective macroprudential stress tests. It is not a systematic research agenda but a set of suggestions for stress testing practitioners. The focus is not on specific models but rather on approaches and practices that can, in the author's view, yield improvements or address obvious pitfalls and, in this way, move the dial on macroprudential stress tests. Needless to say, the priorities these suggestions reflect are—inevitably—subjective, and should be seen as an attempt to contribute to a broader conversation.

Use a variety of models

Given the limitations of the existing stress testing frameworks, it is surprising to see several central banks and regulatory agencies relying on a single modeling approach—typically an expanded balance sheet-based model.[16] This has advantages: these models have been developed and validated in-house (often at a considerable resource investment); are tailored to the characteristics and granularity of the available data; and include features that are particularly well-suited to the type of risks faced by the financial system in question. Clarity and ease of communication of the results—to senior policy-makers or the broader public—are also easier to achieve using a single model. But the preceding discussion has underscored the main risk of this approach: making the outcome of the stress testing exercise hostage to the limitations of a single analytical framework.

---

[16] This, of course, refers to top-down stress tests. Bottom-up stress tests, such as the CCAR of the U.S. Federal Reserve, involve a variety of modeling approaches of the individual banks and benefit from more granular information and a better understanding of idiosyncratic risks.

Instead, a variety of models should be used for macroprudential stress testing, taking the same shock scenario(s) as the starting point. In particular, both balance sheet and market price-based models should be used, to provide different perspectives on risk. This could be done in the context of a single stress testing exercise, like the periodic tests conducted by many central banks in the context of Financial Stability Reports; or over time, with market price-based models providing a vehicle for continuous systemic risk monitoring and balance sheet-based models providing infrequent but more detailed analyses based on actual bank balance sheet data.

The challenge in an approach like this is to interpret and synthesize the results of the different models into a coherent and persuasive narrative. Should the different results be combined or averaged according to a strict rule? Should qualitative judgment be used in weighing different—and potentially contradictory—results? Would this expose the central bank or supervisory agency to the accusation that it is "spinning" the stress test results so as to convey a specific message to the banks or the public? These are not easy questions. But this is a challenge well worth tackling, as it would enhance the insights into systemic risk and the quality of the ensuing conversation about financial stability, both within the supervisory agency and with the banks.

These benefits have been recognized by the IMF, which often combines different analytical models for stress testing in the context of individual Financial Sector Assessment Programs (FSAPs). In its new framework for stress testing the U.K. banking system, the Bank of England also expects to use "a suite of models" to gauge the impact of stress scenarios (Bank of England 2013).

Run more—and better—stress scenarios

Even when a variety of analytical approaches is used—for example, in the case of bottom-up stress tests—the exercise is often limited to one or two macroeconomic stress scenarios (for instance, an "adverse" and a "severe" scenario, in addition to the baseline). But as the discussion of the robustness problem has shown, this approach has a major pitfall: resilience to a shock of a given probability does not imply resilience to *all* shocks with the same probability. From this point of view, the outcome of a test of a single stress scenario may be misleading: if it suggests that the system remains robust under a severe shock, it may convey a false sense of security; if, on the contrary, it shows large systemic losses, it may be unduly alarmist.

Absent an analytically rigorous approach of estimating maximum systemic losses for a multidimensional region with a given probability mass—as discussed in the previous section—the obvious second best is to use a multitude of (extreme but plausible) scenarios for the stress tests and report a distribution of results around the main stress scenario

forecast.[17] This would provide a better sense of the resilience of the system and of individual institutions to a range of shocks than a single scenario.

Using multiple scenarios (as well as a variety of models) would also have another big advantage: it would minimize the scope for individual institutions to "game the test," i.e., make portfolio choices geared toward passing the specific stress test—a risk that was recognized early on (Office of Financial Research 2012; Bank of England 2013).

So why do major central banks and supervisory agencies limit the number of scenarios used in stress tests, often to just one or two? To put it simply—and perhaps a little uncharitably— the answer is cost and convenience. Just like the use of a variety of analytical approaches, using multiple scenarios would make synthesizing and communicating the results of the stress tests much more challenging. In addition, where these tests are bottom-up, as in the U.S.A. and the EU, there is a limit on how many resources the banks can be asked to devote to them.

In addition to the *number*, a related issue is the *type* of scenarios used in stress tests. In most cases, the main stress scenario is an adverse macroeconomic shock, like the "recession scenario" used in the Dodd-Frank stress tests in the U.S.A. The tests essentially postulate a severe macroeconomic shock exogenous to the financial sector and pose the question: how deep and/or prolonged an economic recession would financial institutions be able to weather while still remaining solvent? The size of the exogenous macroeconomic shock is, in turn, typically calibrated on historical experience.[18]

While answering this question would no doubt convey useful information about the resilience of the system (leaving aside for a moment the point about the number of scenarios), is this the right question? The conceptual premise is that negative shocks to domestic macroeconomic factors drive stress events in the financial system. This is certainly true in a number of cases. But in many other cases, the shock originates entirely inside the financial system and is then followed by a recession; and in still others, risk has accumulated inside the financial system to such a degree that a relatively mild exogenous (macro or asset price) shock tips the system "over the edge" and triggers a full-blown financial crisis. In a well-known study of 43 banking crises in 30 countries, Alfaro and Drehmann show that only about half were preceded by adverse macroeconomic conditions and, furthermore, that in

---

[17] See Breuer 2008 and 2010 and Glasserman et al. 2015 for a review of the progress in modeling stress scenarios.

[18] The U.S. Federal Reserve has provided an insight into its decision to calibrate the stress scenario for the Dodd-Frank tests on historical experience rather than on an (analytically more correct) probabilistic approach, i.e., choosing a stress scenario of a certain probability level. Two reasons were given. First, "adverse economic outcomes are typically thought of in terms of how variables evolve in an absolute sense rather than how far away [from the baseline] they lie in the probability space;" therefore, "a scenario featuring a recession may be somewhat clearer and more straightforward to communicate." Second, "the probabilistic approach relies on estimates of uncertainty around the baseline scenario and such estimates are in practice model-dependent" (Federal Reserve System 2013).

nearly 70 percent of all cases in their sample, stress scenarios calibrated on past history fell short of the severity of actual events (Alfaro and Drehmann 2009).

An additional—and increasingly important—aspect of the scenarios that should be used for macroprudential stress tests is their international context. Banks and other financial institutions that may be covered by the tests are increasingly interlinked across borders, and some are truly global. These institutions may be vulnerable to shocks that originate in—or propagate through—a foreign country or market. This requires national central banks or supervisory authorities to design stress test scenarios that explicitly incorporate regional or global dimensions in an internally coherent manner. It may also require closer coordination between stress testers in different countries toward harmonizing scenario design and calibration, data requirements, etc. This is an area that is only now starting to attract the attention of regulators (see Bank of England 2013).

On the basis of this discussion, effective macroprudential stress tests should involve a higher number and a wider range of "smart" stress scenarios, in addition to the one or two standard macroeconomic recession scenarios typically used. Clearly, this would entail a number of challenges: it would require an in-depth understanding of the risk factors affecting the financial system, including cross-border dimensions, so as to tailor stress scenarios to those factors, rather than simply assume a generalized macro recession; it would complicate the task of synthesizing and communicating the results—especially when accompanied by a suite of analytical approaches; and it would involve higher costs. It is these challenges that have held back many supervisors from moving in this direction. However, given the significant pitfalls of limiting the number of scenarios to just one or two, it may be time to reconsider the cost-benefit balance of the current approach.

Expand coverage to non-bank financial entities

Microprudential stress tests have been traditionally applied to banks for two main reasons. First, banks were historically the predominant agents of financial intermediation, and it has long been understood that intermediating retail deposits, especially when these are insured, involves information asymmetries and moral hazard. These, in turn, imply that banks, left to their own devices, will tend to take excessive risks and not deliver the socially optimal amount of financial stability. Second, and relatedly, the machinery of prudential oversight means that a wealth of granular data about the risks facing banks is available to supervisors, who can use them to implement stress testing models. Other financial entities, like mutual funds, asset managers, leasing, etc. either do not engage directly in deposit intermediation or let investors bear all the risks or, in some cases, are too small to be systemically important. Therefore, they do not merit the same regulatory burden as banks.

Recent trends have undermined this received wisdom. For a number of reasons, the line between banks and non-banks has been blurred: banks now engage in what have traditionally been nonbank activities (e.g., investment banking); and non-banks are increasingly undertaking quasi-intermediation functions (shadow banking). Moreover, the non-banking industry has expanded greatly in size and importance in the last two decades (IMF 2015), in part due to the lighter regulatory burden compared to banks. Last but not least, the global

financial crisis has amply demonstrated that banks and non-banks are deeply interconnected, and risks move easily between the two. Therefore, from a macroprudential perspective, systemic risk assessment cannot ignore the non-banking sectors.

Regulators have responded with so-called "structural measures" (the proposals by Volcker, Vickers, and Liikanen that aim at re-drawing a sharp line between bank and non-bank activities), as well as with greater regulatory scrutiny on non-banks. The latter includes stress testing for insurance companies and, more recently, central counterparties (CCPs). For the most part, however, these stress tests are conducted separately from those on banks. While this helps ensure the soundness of individual non-bank entities, these stress tests do little to address the need for a proper systemic risk assessment.

Market price-based models are ideal for integrating banks and non-banks into a single stress testing framework and accounting for risk correlations among them. Balance sheet-based stress testing frameworks can use a network "overlay" that includes non-banks to explore the possibility of contagion through direct exposures across sectors; but as discussed earlier, this is not sufficient to capture all possible channels of contagion. Market price-based models, such as the distress dependence model or the systemic CCA, on the other hand, can easily incorporate non-banks, especially (though not exclusively) entities for which market data are available to estimate individual default probabilities. This flexibility, together with their ability to perform high-frequency tests, is perhaps the most compelling argument for making market price-based models a standard feature of any macroprudential stress testing framework.

The choice of what non-bank entities to incorporate into the stress testing framework would depend on country circumstances. Priority should be given to sectors that are closely connected with banks through ownership or financial linkages, typically insurance companies. Asset management companies, mutual funds, and sometimes pension funds are also sometimes important providers of liquidity to banks, and could thus be affected by—or be a propagation channel for—a systemic shock.

Explore agent-based models

Microprudential stress tests are focused on the ability of a single institution to remain solvent and liquid and continue to fulfill regulatory requirements under stress. Macroprudential stress tests, on the other hand, are supposed to focus on the resilience of the system as a whole and capture general equilibrium dimensions of financial stress, including in particular the behavioral responses of individual institutions to shocks and the possibility of contagion through interconnectedness.

The previous sections have explored both the successes and the pitfalls of the efforts to move from the former to the latter. Most of these efforts have involved starting with well-established (balance sheet or market price-based) models focused on individual banks, and expanding or adapting them in various ways for macroprudential purposes.

But all these models—like all traditional economic and financial models based on neoclassical micro foundations—face a more fundamental problem: they assume homogeneous agents (individuals or institutions) that always behave rationally in ways that can be modeled based on past experience (during normal, non-crisis times); and that policy decisions influence this behavior in the same way for all market participants. These assumptions miss some critical points about financial crises,[19] notably:

- the fact that that market participants are heterogeneous and often make less-than-rational decisions, especially under stress;

- the emergence of a new dynamic under stress, when relationships among financial institutions can change in diverse and complex ways (the endogenous network literature, for instance, has explored how network formation changes depending on the environment—see Deb 2015);

- the fact that the response of regulated institutions to policy signals is state-contingent (for example, Klinger and Teplý 2014 show how raising the regulatory capital requirements put in place in normal times to ensure banks have sufficient capital buffers has almost no positive effect on systemic stability at times of crisis).

Agent-based models can capture many of these aspects that traditional models miss, and are in principle better suited for analyzing situations of financial stress, as argued strongly by Bookstaber (2012). An agent-based model postulates autonomous, heterogeneous agents with bounded rationality, and specifies heuristic rules that dictate how they will act based on various factors. These rules can vary across different types of agents (for instance banks, depositors, providers of wholesale funding) and allow for less-than-optimal behavior. The model determines the "topology," that is the mechanism through which agents can interact (for example, how they form networks), and can explore various types of shocks, both exogenous and endogenous (such as changes in agent behavior, topology rules, etc.). Agent-based models are implemented through the application of software code, typically for Monte Carlo simulations.

---

[19] It has been argued that these models, exemplified by the Dynamic Stochastic General Equilibrium (DGSE) framework widely used today, miss critical aspects of economic behavior *even in normal times*. Robert Solow has this to say: "[These models] take it for granted that the whole economy can be thought about as if it were a single, consistent person or dynasty carrying out a rationally designed, long-term plan, occasionally disturbed by unexpected shocks, but adapting to them in a rational, consistent way. […] The DSGE School populates its simplified economy—remember that all economics is about simplified economies just as biology is about simplified cells—with exactly *one* single combination worker-owner-consumer-everything-else who plans ahead carefully and lives forever. One important consequence of this 'representative agent' assumption is that there are no conflicts of interest, no incompatible expectations, no deceptions. […] This cannot be an adequate description of a national economy, which is pretty conspicuously *not* pursuing a consistent goal. A thoughtful person, faced with the thought that economic policy was being pursued on this basis, might reasonably wonder what planet he or she is on" (Solow 2010). Regardless of how valid one considers Solow's criticism, it is clearly much more relevant at times of macroeconomic or financial stress, which is what stress testing frameworks— the narrow focus of this paper—are supposed to capture.

Agent-based models are increasingly being used for macrofinancial modeling. A well-known example is the Complexity Research Initiative for Systemic Instabilities (CRISIS), building a large-scale macrofinancial agent-based model for the European economy.[20] Relatively simpler versions are used by Klinger and Teplý (2014) and Chan-Lau (2014) to explore the impact of financial regulation on bank solvency, liquidity, and contagion under stress scenarios (the latter paper also contains a brief literature survey on the use of agent-based models for macrofinancial modeling).

Agent-based models are complex, and implementing them would require a shift in the approaches traditionally taken by (and the skills traditionally required of) stress testers. Nevertheless, the limited experience so far suggests that they can provide unique insights into the aspects that matter most in a stress scenario: the behavioral responses of banks and the interactions between banks, market participants, and policy-makers. In addition, unlike market price-based models, agent-based model simulations provide results that are tractable and allow for the emergence of outcomes that could not have been predicted based on the past behavior of individual agents in the financial system. Therefore, they are very promising tools, and there is a strong argument for expanding their use in stress testing.

Embed stress tests into the financial stability policy framework

This injunction may seem surprising: whatever their shortcomings, stress tests have always been part of the toolkit for the analysis of financial vulnerabilities, and their results inform policy decisions targeted at restoring or maintaining the stability of individual financial institutions and, increasingly, the system as a whole.

The point, however, is that the recent explosion of interest in stress testing is creating a risk. Policy-makers, market participants, and the broader public may focus excessive attention on stress tests, form exaggerated expectations, take stress test results out of context, and give them much greater weight than they merit in guiding policy action. This risk is evident in the way stress test results tend to dominate the public debate on the health of the banks in the U.S.A. following the Dodd-Frank Act, as well as in Europe following a string of highly-publicized tests by the European Banking Authority. Immediate remedial action by bank management and/or supervisors is automatically expected—or indeed required—of banks "failing" the test. This is a commendable demonstration of resolve on behalf of policy-makers (the alternative—conducting high-profile stress tests and failing to take any action after the results—would almost certainly be worse). It can also be seen as a resounding vote of confidence in the power of stress tests to assess correctly the resilience of financial institutions, as well as. But after reading about the pitfalls of the current stress testing frameworks, especially as regards their ability to assess systemic risk, the reader may well ask: is this confidence justified?

---

[20] CRISIS is a multi-year project involving a consortium of universities, private firms, and public entities, launched in 2011 and financed by the EU. For more information, see http://www.crisis-economics.eu/.

This was less of an issue with the traditional microprudential stress tests. Those tests were an input into the assessment of soundness of individual institutions. Their results were not made public and rarely triggered automatic remedial action by themselves; they were instead used to inform the ongoing conversation between the regulated entity and the regulator. But the unprecedented attention focused on recent system-wide stress tests in advanced economies seems at times to overshadow, rather than inform, the conversation about financial stability among policy-makers, regulators, individual institutions, market participants, and the public.

This risk has been noted before. In setting out best-practice principles for macroprudential stress testing, the IMF put it this way (IMF 2012):

> "Regardless of how extensive the coverage of risk factors, how refined the analytical models, how severe the shocks incorporated in the stress tests, and how careful the communications strategy, there is always the risk that the 'unthinkable' will materialize. […] No matter how hard the stress tester tries, stress tests always have margins of error. Their results will almost always turn out to be optimistic or pessimistic ex post. In addition, there will always be model risk, imperfect data, or underestimation of the severity of the shock. One should therefore set stress test results in a broader context."

Contrast this with the approach in the Bank of England's Discussion Paper on the framework for stress testing the UK banking system. In this document, one of the purposes of regular stress tests is to "be a device through which the Bank [of England] can be held accountable to Parliament, and the wider public, on its financial stability objective; […] a tool through which the Financial Policy Committee and the Prudential Regulation Authority Board can articulate the resilience standard against which they aim to hold the system" (Bank of England 2013, p. 13). A literal reading of this could be seen as suggesting that the macroprudential stress tests can provide a well-defined standard of systemic resilience; that if the system "passes" the hurdle rate in the hypothetical stress scenario—or if institutions that fail take the required remedial action—it is guaranteed to be resilient to all real-life shocks of equal or lesser severity; that the regulator stands behind this guarantee; and that in case of a systemic disruption from any shock that is milder than that assumed in the last stress test, the regulator can be held accountable. Such a literal reading is a caricature of what the authors had in mind. Indeed, in a different passage, the Discussion Paper makes clear that "no single quantitative output of the stress testing exercise is expected to be linked mechanically to policy responses." But this caricature highlights the risk of creating expectations that macroprudential stress tests, at least at their current state, would be hard pressed to meet.

So the call to embed stress tests firmly in the financial stability framework is essentially a call for caution and humility. Macroprudential stress testing is just one of the many tools available to assess systemic vulnerabilities and resilience. They should be treated as complements to other tools, such as early warning indicators, and—crucially—should be combined with microprudential perspectives. They should be tailored to the business models and main risk transmission channels in the system, rather than rely on generic models and assumptions. And final conclusions about the resilience of the system to shocks—and about policies to enhance it—should draw on all these sources, not just the results of the stress tests.

REFERENCES

Acharya, V. V. and T. Yorulmazer, 2008, "Cash-in-the-Market Pricing and Optimal Resolution of Bank Failures," *Review of Financial Studies*, Vol. 21, pp. 2705–2742.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson, 2010, "Measuring Systemic Risk," Working Paper No. 10/02 (Cleveland: Federal Reserve Bank).

Acharya, V. V., R. Engle, and D. Pierret, 2013, "Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights," NBER Working Paper No. 18968 (Cambridge, MA: National Bureau of Economic Research).

Adrian, T. and M. K. Brunnermeier, 2011, "COVAR," NBER Working Paper No. 17454 (Cambridge, MA: National Bureau of Economic Research).

Afonso, G., A. Kovner, and A. Shoar, 2010, "Stressed, Not Frozen: The Federal Funds Market in the Financial Crisis," Federal Reserve Bank of New York Staff Report No. 437 (New York: Federal Reserve Bank).

Aikman, D., P. Alessandri, B. Eklund, P. Gai, S. Kapadia, E. Martin, N. Mora, G. Sterne, and M. Willison, 2009, "Funding Liquidity Risk in a Quantitative Model of Systemic Stability," Bank of England Working Paper No. 372 (London: Bank of England).

Alessandri, P., P. Gai, S. Kapadia, N. Mora, and C. Puhr, 2009, "Towards a Framework for Quantifying Systemic Stability," *International Journal of Central Banking*, Vol. 5, No. 3, pp. 47–81.

Allen, F. and D. Gale, 2007, *Understanding Financial Crises* (New York: Oxford University Press).

Alfaro, R. and M. Drehmann, 2009, "Macro Stress Tests and Crises: What Can We Learn?" *BIS Quarterly Review,* December.

Altman, E. I., 1968, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance,* Vol. 23, No. 4, pp. 589–609.

Altman, E. I. and E. Hotchkiss, 2006, *Corporate Financial Distress and Bankruptcy*, 3rd edition, (New Jersey: John Wiley & Sons, Inc.).

Anand, K., G. Bédard-Pagé, and V. Traclet, 2014, "Stress Testing the Canadian Banking System: A System-Wide Approach," *Financial System Review,* June, pp. 61–68.

Aymanns, C., C. Caceres, C. Daniel, and L. Schumacher, 2015, "Bank Solvency and Funding Cost," IMF Working Paper (forthcoming), (Washington DC: International Monetary Fund).

Bank of England, 2013, "A Framework for Stress Testing the UK Banking System," Bank of England Discussion Paper, October (London: Bank of England).

Bank of Korea, 2012, "Systemic Risk Assessment Model for Macroprudential Policy (SAMP)," *Financial Stability Report*, October (Seoul: Bank of Korea).

Barnhill, T. and L. Schumacher, 2011, "Modeling Correlated Systemic Liquidity and Solvency Risks in a Financial Environment with Incomplete Information," IMF Working Paper 11/263 (Washington DC: International Monetary Fund).

Basel Committee on Banking Supervision, 2005, "International Convergence of Capital Measurement and Capital Standards: A Revised Framework," Technical Report (Basel: Bank for International Settlements).

Basel Committee on Banking Supervision, 2009, *Revisions to the Basel II Risk Framework* (Basel: Bank for International Settlements).

Basel Committee on Banking Supervision, 2013, "Liquidity Stress Testing: A Survey of Theory, Empirics and Current Industry and Supervisory Practices," Basel Committee on Banking Supervision Working Paper No. 24 (Basel: Bank for International Settlements).

Bernanke, B. S., 2008, "Reducing Systemic Risk," remarks at the Federal Reserve Bank of Kansas City's Annual Economic Symposium, Jackson Hole, WY (Washington DC: Federal Reserve).

Bernanke, B. S., 2013, "Stress Testing Banks: What Have We Learned?" remarks at "Maintaining Financial Stability: Holding a Tiger by the Tail," a Financial Markets Conference sponsored by the Federal Reserve Bank of Atlanta (Washington DC: Federal Reserve).

Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol. 81, No. 3, pp. 637-654.

Board of Governors of the Federal Reserve System, 2014, *Comprehensive Capital Analysis and Review 2015: Summary Instructions and Guidance* (Washington DC: Federal Reserve).

Bolton, P., T. Santos, and J. Scheinkman, 2009, "Market and Public Liquidity," *American Economic Review*, Vol. 99, pp. 594–599.

Bookstaber, R., 2012, "Using Agent-Based Models for Analyzing Threats to Financial Stability," Office of Financial Research Working Paper No. 0003 (Washington DC: U.S. Department of the Treasury).

Borio, C., 2003, "Towards a Macroprudential Framework for Financial Supervision and Regulation?" BIS Working Paper No. 128 (Basel: Bank for International Settlements).

Borio, C., M. Drehmann, and K. Tsatsaronis, 2012, "Stress Testing Macro Stress Testing: Does it Live Up to Expectations?" BIS Working Paper No. 369 (Basel: Bank for International Settlements).

Breuer, T., 2008, "Overcoming Dimensional Dependence of Worst Case Scenarios and Maximum Loss," *Journal of Risk*, Vol. 11, No. 1, pp. 79–92.

Breuer, T., M. Jandačka, K. Rheinberger, and M. Summer, 2009, "How to Fund Plausible, Severe, and Useful Stress Scenarios," *International Journal of Central Banking*, Vol. 5, No. 3, pp. 205–224.

Chan-Lau, J.A., 2010, "Regulatory Capital Charges for Too-Connected to Fail Institutions," *Financial Markets, Institutions, and Instruments*, Vol. 19, No. 5, pp. 355 – 79.

Chan-Lau, J. A., 2013, *Systemic Risk Assessment and Oversight*, (London: Risk Books).

Chan-Lau, J. A., 2014, "Regulatory Requirements and Their Implications for Bank Solvency, Liquidity, and Interconnectedness Risks: Insights from Agent-Based Model Simulations." Available at SSRN: http://ssrn.com/abstract=2537124.

Crockett, A. D., 2000, "Marrying the Micro- and Macro-Prudential Dimensions of Financial Stability," remarks before the Eleventh International Conference of Banking Supervisors, Basel, 20-21 September 2000 (Basel: Bank for International Settlements).

Crosbie, P. and J. Bohn, 2003, *Modeling Default Risk: Modeling Methodology* (New York: Moody's KMV).

Deb, P., 2015, "Market Frictions, Interbank Linkages, and Excessive Interconnections," IMF Working Paper (forthcoming) (Washington DC: International Monetary Fund).

Diamond, D. W. and R. G. Rajan, 2011, "Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes," *Quarterly Journal of Economics*, Vol. CXXVI, issue 2, pp. 557–591.

Dwyer, D. W., A. E. Kocagil, and R. S. Stein, 2004, *Moody's KMV Riskcalc™ v3.1 Model* (New York: Moody's KMV).

European Banking Authority, 2014, *Methodological Note EU-Wide Stress Test 2014*. Available at: https://www.eba.europa.eu/documents/10180/669262/Methodological+Note.pdf.

Federal Reserve System, 2013, "Policy Statement on the Scenario Design Framework for Stress Testing," *Federal Register*, Vol. 78, No. 230, pp. 71435–71448.

Feldkircher, M., G. Fenz, R. Ferstl, G. Krenn, B. Neudorfer, C. Puhr, T. Reininger, S. W. Schmitz, M. Schneider, C. Siebenbrunner, M. Sigmund, and R. Spitzer, 2013, "ARNIE in Action: the 2013 FSAP Stress Tests for the Austrian Banking System," *Financial Stability Report,* No. 26, pp. 100–118.

Gauthier, C., A. Lehar, and M. Souissi, 2010, "Macroprudential Regulation and Systemic Capital Requirements," Bank of Canada Working Paper No. 2010–4 (Ottawa: Bank of Canada).

Glasserman, P., C. Kang, and W. Kang, 2015, "Stress Scenario Selection by Empirical Likelihood," *Quantitative Finance*, Vol. 15, no. 1, pp. 25-41.

Goodhart, C., P. Sunirand, and D. Tsomocos, 2006, "A Model to Analyze Financial Fragility," *Economic Theory*, Vol. 27, No. 1, pp. 107–142.

Gorton, G. and A. Metrick, 2009, "Securitized Banking and the Run on the Repo," NBER Working Paper No. 15223 (Cambridge, MA: National Bureau of Economic Research).

Gray, D. F. and S. W. Malone, 2008, *Macrofinancial Risk Analysis* (Hoboken: John Wiley & Sons, Inc.).

Gray, D. F., 2010, "New Directions in Financial Sector and Sovereign Risk Management," *Journal of Investment Management*, Vol. 8, No. 1, pp. 23–38.

Gray, D. F. and A. A. Jobst, 2011, "Modelling Systemic Financial Sector and Sovereign Risk," *Sveriges Riksbank Economic Review*, No. 2, pp. 68–106.

Greenlaw, D., Kashyap, A. K., Schoenholtz, K., and H. S. Shin, 2012, "Stressed Out: Macroprudential Principles for Stress Testing," Booth School of Business Working Paper No. 71(Chicago: The University of Chicago).

Heider, F., M. Hoerova, and C. Holthausen, 2009, "Liquidity Hoarding and Interbank Market Spreads: The Role of Counterparty Risk," ECB Working Paper No. 1126 (Frankfurt: European Central Bank).

Henry, J., C. Kok, A. Amzallag, P. Baudino, I. Cabral, M. Grodzicki, M. Gross, G. Halaj, M. Kolb, M. Leber, C. Pancaro, M. Sydow, A. Voulidis, M. Zimmermann, and A. Zochowski, 2013, "A Macro Stress testing Framework for Assessing Systemic Risks in the Banking Sector," ECB Occasional Paper No. 152 (Frankfurt: European Central Bank).

Huang, X., H. Zhou, and H. Zhu, 2009, "A Framework for Assessing the Systemic Risk of Major Financial Institutions," *Journal of Banking and Finance*, Vol. 33, no. 11, pp. 2036 – 49.

International Monetary Fund, 2008, *Global Financial Stability Report, October 2008* (Washington DC: International Monetary Fund).

International Monetary Fund, 2010, "United States: Financial Sector Assessment Program Documentation—Technical Note on Stress Testing," IMF Country Report No. 10/244. Available at http://www.imf.org/external/pubs/ft/scr/2010/cr10244.pdf.

International Monetary Fund, 2011a, "Sweden: Technical Note on Contingent Claims Analysis to Measure Risk and Stress Test the Swedish Banking Sector," IMF Country Report No. 11/286. Available at http://www.imf.org/external/pubs/ft/scr/2011/cr11286.pdf.

International Monetary Fund, 2011b, "United Kingdom: Technical Note on Stress Testing the Banking Sector," IMF Country Report No. 11/227. Available at http://www.imf.org/external/pubs/ft/scr/2011/cr11227.pdf.

International Monetary Fund, 2012a, *Macrofinancial Stress Testing: Principles and Practices*, IMF Policy Paper. Available at http://www.imf.org/external/np/pp/eng/2012/082212.pdf.

International Monetary Fund 2012b, "Israel: Technical Note on Stress testing the Banking, Insurance and Pension Sectors," IMF Country Report No. 12/88. Available at http://www.imf.org/external/pubs/ft/scr/2012/cr1288.pdf.

International Monetary Fund, 2014a, "Switzerland: Financial Sector Stability Assessment," IMF Country Report No. 14/143. Available at http://www.imf.org/external/pubs/ft/scr/2014/cr14143.pdf.

International Monetary Fund, 2014b, "People's Republic of China—Hong Kong Special Administrative Region: Technical Note on Stress Testing the Banking Sector," IMF Country Report No. 14/210. Available at http://www.imf.org/external/pubs/ft/scr/2014/cr14210.pdf.

International Monetary Fund, 2015, *Global Financial Stability Report, April 2015* (Washington DC: International Monetary Fund).

Jobst, A. A. and D. F. Gray, 2013, "Systemic Contingent Claims Analysis—Estimating Market-Implied Systemic Risk," IMF Working Paper 13/54 (Washington DC: International Monetary Fund).

Kapadia, S., M. Drehmann, J. Elliott, and G. Sterne, 2012, "Liquidity Risk, Cash-Flow Constraints and Systemic Feedbacks," Bank of England Working Paper No. 456 (London: Bank of England).

Kitamura, T., S. Kojima, K. Nakamura, K. Takahashi, and I. Takei, 2014, "Macro Stress Testing at the Bank of Japan," Bank of Japan Reports & Research Papers. Available at: http://www.boj.or.jp/en/research/brp/ron_2014/ron141008a.htm/.

Klinger, T. and P. Teplý, 2014, "Systemic Risk of the Global Banking System—An Agent-Based Network Model Approach," *Prague Economic Papers*, Vol. 1, pp. 24–41.

Merton, R. C., 1974, "On the Pricing of Corporate Debt: the Risk Structure of Interest Rates," *Journal of Finance*, Vol. 29, No. 2, pp. 449–470.

Neyman, A., 2002, "Value of Games With Infinitely Many Players," in Aumann, R.J. and S. Hart (eds.), *Handbook of Game Theory with Economic Applications* (Amsterdam: Elsevier).

Office of Financial Research, 2012, *Annual Report (*Washington DC: U.S. Department of the Treasury).

Pierret, D., 2014, "Systemic Risk and the Solvency-Liquidity Nexus of Banks," mimeo, NYU Stern School of Business.

Pritsker, M., 2014, "Enhanced Stress Testing and Financial Stability," mimeo, The Federal Reserve Bank of Boston.

Puhr, C. and S. W. Schmitz, 2013, "A View from the Top—The Interaction Between Solvency and Liquidity Stress," *Journal of Risk Management in Financial Institutions*, Vol. 7, No. 1, pp. 38–51.

Segoviano, M. A., 2006, "Consistent Information Multivariate Density Methodology," Financial Markets Group, London School of Economics Discussion Paper No. 557 (London: London School of Economics).

Segoviano, M. A. and C. Goodhart, 2009, "Banking Stability Measures," IMF Working Paper 09/4 (Washington DC: International Monetary Fund).

Segoviano, M. A. and C. Goodhart, 2010, "Distress Dependence and Financial Stability," in *Financial Stability, Monetary Policy, and Central Banking*, ed. by Alfaro (Santiago: Central Bank of Chile).

Segoviano, M. A., S. Malik, P. Lindner, and F. Cortes, 2015, "Systemic Risk and Interconnectedness (SyRin): A Comprehensive Multi-Sector Framework," IMF Working Paper (forthcoming), (Washington DC: International Monetary Fund).

Solow, R., 2010, "Building a Science of Economics for the Real World," in *Statement at the Hearing of the House Committee on Science and Technology Subcommittee on Investigations and Oversight*, U.S. House of Representatives, 111th Congress, July 20 (Washington DC: U.S. Government Printing Office).

Stein, J. and A. Kashyap, 2004, "Cyclical Implications of the Basel II Capital Standards," *Economic Perspectives*, Vol. 28, Q1, pp. 18–31.

Tarashev, N., C. Borio, and K. Tsatsaronis, 2010, "Attributing Systemic Risk to Individual Institutions," BIS Working Paper No. 308 (Basel: Bank for International Settlements).

Tarullo, D. K., 2014, "Stress Testing after Five Years," remarks at the Federal Reserve Third Annual Stress Test Modeling Symposium (Boston: Federal Reserve Bank).

Webber, L. and M. Willison, 2011, "Systemic Capital Requirements," Bank of England Working Paper No. 436 (London: Bank of England).

Wong, E. and C. Hui, 2009, "A Liquidity Stress-Testing Framework with Interaction between Market and Credit Risks," Hong Kong Monetary Authority Working Paper No. 06/2009 (Hong Kong: Hong Kong Monetary Authority).

Zangari, P., 1996, "Statistical and Probability Foundations," *RiskMetricTM—Technical Document*, 4th edition, Ch. 4, pp. 45–74 (New York: Morgan Guaranty Trust Company of New York).