Working Paper

INTERNATIONAL MONETARY FUND

WP/07/59

IMF Working Paper

Introduction to Applied Stress Testing

Martin Čihák

INTERNATIONAL MONETARY FUND

 WP/07/59

IMF Working Paper

Monetary and Capital Markets Department

Introduction to Applied Stress Testing

Prepared by Martin Čihák[1]

Authorized for distribution by Mark Swinburne

March 2007

Abstract

Stress testing is a useful and increasingly popular, yet sometimes misunderstood, method of analyzing the resilience of financial systems to adverse events. This paper aims to help demystify stress tests, and illustrate their strengths and weaknesses. Using an Excel-based exercise with institution-by-institution data, readers are walked through stress testing for credit risk, interest rate and exchange rate risks, liquidity risk and contagion risk, and are guided in the design of stress testing scenarios. The paper also describes the links between stress testing and other analytical tools, such as financial soundness indicators and supervisory early warning systems. Furthermore, it includes surveys of stress testing practices in central banks and the IMF.

JEL Classification Numbers: G10, G20
Keywords: stress testing, financial soundness indicators, early warning systems
Author's E-Mail Address: mcihak@imf.org

[1] I would like to thank R. Sean Craig, Dale Gray, Plamen Iossifov, Peter Chunnan Liao, Thomas Lutton, Christiane Nickel, Nada Oulidi, Richard Podpiera, Leah Sahely, Graham Slack, participants in a regional conference on financial stability issues at Sinaia, Romania, and in seminars at the IMF, the World Bank, the Central Bank of Russia, and the Central Bank of Trinidad and Tobago for helpful comments. I would also like to thank Matthew Jones and Miguel Segoviano for insights on stress testing methodologies, and Roland Straub for help with the overview of stress tests in Appendix II. All remaining errors are mine.

Contents Page

Table

Figures

Boxes

Appendixes

Appendix Tables

I. Introduction

In response to increased financial instability in many countries in the 1990s, policy makers, researchers, and practitioners became interested in better understanding vulnerabilities in financial systems (e.g., Crockett, 1997). One of the key techniques for quantifying vulnerabilities is stress testing.

Stress testing is a general term encompassing various techniques for assessing resilience to extreme events. Stress tests are used to determine the stability of a given system or entity. They involve testing beyond normal operational capacity, often to a breaking point, in order to observe the results. In financial literature, stress testing has traditionally referred to asset portfolios, but more recently it has been applied to whole banks, banking systems, and financial systems.[2]

The subject of this paper and the accompanying Excel file are system-oriented stress tests carried out on bank-by-bank data.[3] The paper and the accompanying file aim to illustrate strengths and weaknesses of stress testing, using concrete examples. Readers will familiarize themselves with how common types of stress tests can be implemented in practice, in a small and non-complex banking system. They should gain an understanding of how the various potential shocks can be fitted together and how stress testing complements other analytical tools, such as financial soundness indicators (FSIs) and supervisory early warning systems. They should also learn how to interpret the results of stress tests.

As the title indicates, this paper is about *applying* stress tests to actual data. It devotes relatively little space to general discussions of what stress testing is. There is already a plethora of studies that provide a general introduction to stress testing, discussing its nature and purpose. For example, Blaschke and others (2001), Jones, Hilbers, and Slack (2004), IMF and World Bank (2005b), and Čihák (2004a, 2005) provide a general introduction to stress testing. In contrast, relatively little is available in terms of practical technical guidance on how to actually implement stress tests for financial systems, using concrete data as examples. This document and the accompanying Excel file are an attempt to fill that gap.

As the title also suggests, this is an *introduction*, not a comprehensive "stress testing cookbook." The paper covers basic versions of the most common stress tests.[4] Depending on

---

[2] For a survey of stress tests done by banks on individual portfolios, see, e.g., CGFS (2005). Outside of finance, the term "stress testing" is used in areas as diverse as cardiology, engineering, and software programming.

[3] We concentrate on banks, even though the impact of risks in non-bank financial institutions is also discussed.

[4] For this reason, the courses or seminars based on this document and the accompanying file have been called "Stress Testing 101" (I have also considered "Stress Testing for Dummies," but have not used it for copyright reasons, among other things).

the sophistication of the financial system and the type of its exposures, it may be necessary to elaborate on these basic versions of tests (e.g., by estimating econometrically some relationships that are only assumed in this file), and perhaps include also tests for other risks (e.g., asset price risks or commodity risks) if financial institutions have material exposures to those risks. The accompanying file can be developed in a modular fashion, with additional modules capturing additional risks or elaborating on the existing ones. The final part of this paper provides an overview of the main extensions that could be considered.

One of the key messages of this paper is that assumptions matter, and they particularly matter in stress testing. The paper calls for transparency in presenting stress testing assumptions, and for assessing robustness of results to the assumptions. **To highlight the importance of assumptions in stress testing, this document uses bold letters for references to assumptions used in the accompanying Excel file.** To achieve transparency in the Excel file itself, assumptions are highlighted (in blue and green) and grouped in one worksheet.

This document has 10 sections, 3 appendixes, and an accompanying Excel file. Section II provides an overview of the general issues one needs to address when carrying out stress tests, and describes the design of the accompanying Excel file. It also introduces the general setting of the fictional economy described by the Excel file. Section III discusses the input data. Sections IV–VIII discuss the stress tests for the individual risk factors, namely credit risk (Section IV), interest rate risk (Section V), foreign exchange risk (Section VI), interbank contagion risk (Section VII), and liquidity risk (Section VIII). Section IX shows how to create scenarios from the individual risk factors, how to present the results, and how to link the results to supervisory early warning systems. Section X concludes and discusses possible extensions. Appendix I contains tasks and questions related to the stress tests that could be used in a workshop or seminar based on this document, or that a reader can use when practicing stress tests with the accompanying Excel file. Appendix II contains an overview of stress tests in selected financial stability reports published by central banks. Appendix III gives an overview of stress tests in Financial Sector Assessment Program (FSAP) missions.

The accompanying Excel file, “Stress Tester 2.0.xls,” constitutes an essential part of this document. For each of the concepts introduced in the subsequent sections, we include specific references to the file. To distinguish references to tables in the Excel file from those in this document, the Excel table names start with a capital letter A–H (denoting the order of the spreadsheet) followed by a number (denoting the order of the table within the spreadsheet). For instance, Table A2 denotes the second table in the first Excel spreadsheet. Tables in the text of this paper are denoted simply by a number.

II. Overview of the File and of the Stress Testing Process

Stress testing can be thought of as a process that includes (i) identification of specific vulnerabilities or areas of concern; (ii) construction of a scenario; (iii) mapping the outputs of the scenario into a form that is usable for an analysis of financial institutions' balance sheets and income statements; (iv) performing the numerical analysis, (v) considering any second round effects; and (vi) summarizing and interpreting the results (Jones, Hilbers, and Slack, 2004; IMF and World Bank, 2005b). The aim of this exercise is to illustrate the stages of this process. It will also illustrate that these stages are not necessarily sequential, as some modification or review of each component of the process may be needed as work progresses.

The stress testing exercise is performed on the banking system in a fictional country named Bankistan. Given that confidentiality restrictions do not allow the IMF to pass on individual bank data, we refer to a fictional country rather than an actual one. The exercise is modeled on stress tests conducted in a number of Financial Sector Assessment Program (FSAP) missions, and the input data were created to make the exercise realistic. However, compared to typical FSAP stress tests, the exercise was simplified significantly to make it suitable for a short workshop that would give an overview of the FSAP stress tests (see IMF and World Bank (2003) and Appendix III of this paper). It is a version that could be used for a non-complex banking system. For larger systems characterized by complex financial institutions and markets, more elaborate tests (described only briefly in this document) may be necessary.

To understand how to design stress testing shocks and scenarios, it is important to have a good understanding of the structure of the financial system and the overall environment in which the system operates. Box 1 therefore provides short briefing information on the macroeconomic and macroprudential situation in Bankistan.

A. How To Operate Stress Tester 2.0—A Quick Guide to the Accompanying File

This section provides an introduction to the accompanying Excel file, "Stress Tester 2.0.xls." The operation of the file requires relatively little prior knowledge. Users should be proficient in operating standard Excel files, and should have read this document. Knowledge of intermediate macroeconomics is useful for understanding the linkages between the financial sector and the broader macroeconomic framework.

The file can be viewed as a module belonging to a broader stress testing framework (Figure 1). Such a framework would typically include a model characterizing linkages among key macroeconomic variables, such as GDP, interest rates, the exchange rate, and other variables. Medium-scale macroeconomic models (e.g., those used by a central bank for macroeconomic forecasts) including dozens of estimated or calibrated relationships are often

Box 1. Background Information on Bankistan's Economy and Banking Sector

The economic environment in which banks are operating in Bankistan is challenging, with increasing macroeconomic imbalances, inappropriate macroeconomic policies, and deep uncertainty fueled by political tensions. Real activity is sharply contracting, and inflation has almost doubled to 65 percent.

Unsustainable fiscal imbalances and loose monetary conditions were key to the deteriorating situation in Bankistan. The government deficit more than doubled in 2005, and a sharp increase in central bank financing of the government has significantly accelerated money growth.

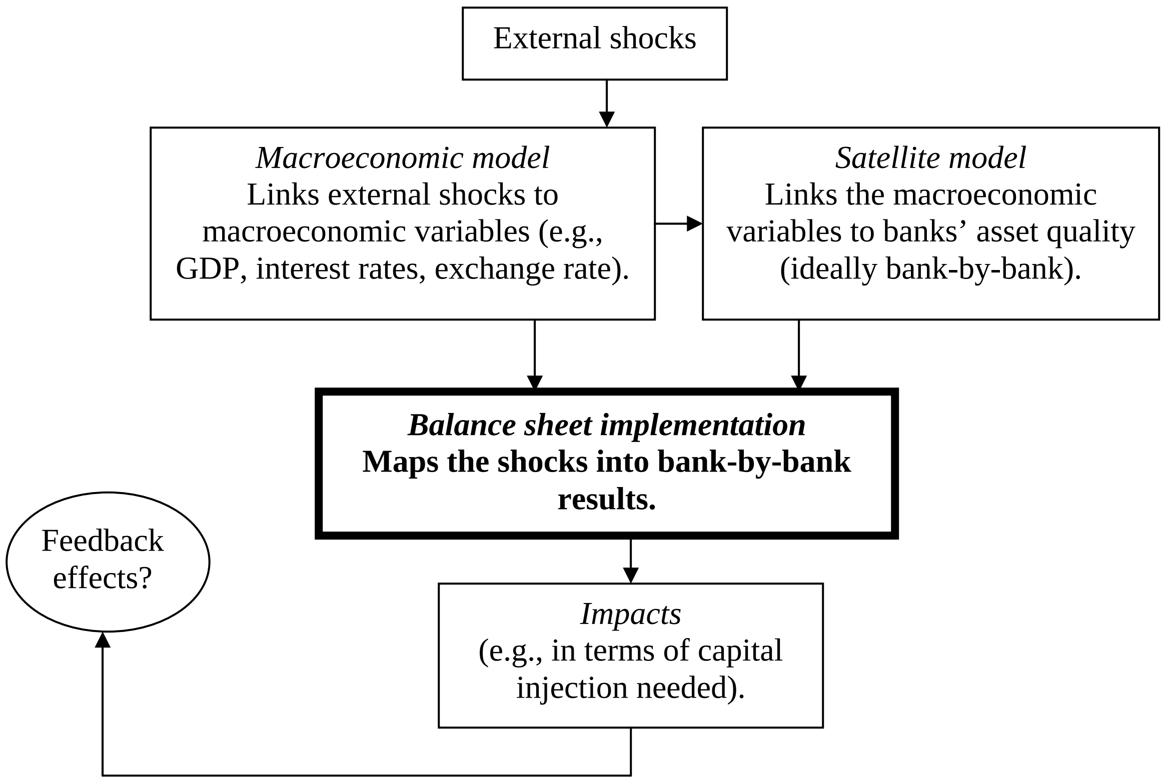
The policy response to the deteriorating situation has been inappropriate. Expansionary monetary policy measures (e.g., a lowering of reserve requirements) have induced a further easing of liquidity conditions. The ensuing excess liquidity induced a drop in treasury bill rates from 60 percent to below 15 percent. This means that together with an inflation rate of 65 percent, real interest rates are sharply negative. (Note: This is used for assessing interest rate risk.)

The official exchange rate of the Bankistan currency, Bankistan dollar (B$), is fixed at 55 B$/US$. However, the black market exchange rate has depreciated in recent months from about 60 B$/US$ to about 85 B$/US$. (Note: This is important information for assessing foreign exchange risk.)

The deteriorating macroeconomic environment has put considerable strain on the financial condition of the banking system. Even though the system has proved so far to be remarkably resilient, some banks have been weakened considerably, and are prone to further deterioration in light of the significant risks. Reported high capital adequacy ratios were found to be overstated due to insufficient provisioning. (Note: This information is used for the assessment of asset quality.) In addition, asset quality has deteriorated. The ratio of gross nonperforming loans (NPLs) to total loans has increased from 15 percent at end-2004 to 20 percent at end-2005. (Note: This information will be used for assessing credit risk.)

The banking system of Bankistan consists of 12 banks. Three of them are state owned (with code names SB1 to SB3), five are domestic privately owned banks (DB1 to DB5), and four are foreign-owned (FB1 to FB4). The banking system, and particularly the state-owned banks, have been plagued by a large stock of NPLs and weak provisioning practices. Data on the structure and performance of the 12 banks are provided in the "Data" sheet of the accompanying Excel file. An assessment of Bankistan's compliance with the Basel Core Principles for Banking Supervision (BCP) suggests that even though existing loan classification and provisioning rules in Bankistan are broadly adequate, they are not well implemented in practice and banks are underprovisioned.

Figure 1. Stress Testing Framework



used for this purpose (if such models are not available, vector autoregression or vector error correction models can be estimated). Given that such models do not generally include financial sector variables, the stress testing framework can also include a "satellite model" that maps (a subset of) the macroeconomic variables into financial sector variables, in particular asset quality. Such a satellite model can be built on data on individual banks over a period of time: using panel data techniques, asset quality in individual banks can be explained as a function of individual bank variables and system-level variables. Together with the macroeconomic model, the satellite model can be used to map assumed external shocks (e.g., a slowdown in world GDP) into bank-by-bank asset quality shocks.

We focus here on how to calculate the bank-by-bank impacts resulting from external shocks, and how to express the impacts in terms of a variable such as capital adequacy or capital injection as a percent of GDP. We spend relatively little time discussing the broader macroeconomic framework, or possible feedback effects to the broader economy. The cells that contain sizes of shocks and numerical assumptions in this file can be thought of as interfaces between this module and the other modules of the stress testing framework. The

cells that contain results (e.g., capital injections as a percent of GDP) can be viewed as interfaces to a module analyzing the feedback effects.

The modular design has the advantage that as a user becomes more experienced in stress testing, or as more data become available for analysis, additional modules can be added or developed. Incorporating, for example, the underlying econometric calculations in Excel would make the resulting file too big and unwieldy (it may also not be possible because the econometric tools in Excel are more limited than in other packages).

All data in the file relate to end-2005, unless indicated otherwise, and are expressed in millions of Bankistan dollars (B$), except for ratios (shown in percent).

The file contains the following nine worksheets: Read Me, Data, Assumptions, Credit Risk, Interest Risk, FX Risk, Interbank, Liquidity, and Scenarios. Table 1 contains a description of the individual worksheets.

To differentiate the various types of cells (input data, numerical assumptions, and formulas), different colors are used in the file. The following color coding is used:

- *Yellow* denotes data reported by the National Bank of Bankistan (NBB). The yellow cells are found only in the “Data” worksheet. When a new set of data arrives from the NBB, the contents of the yellow cells should be replaced with the new data, and all the results are recalculated automatically.
- *Blue* denotes the assumed sizes of the shocks to risk factors, e.g., an increase in interest rates. The blue cells are found only in the Assumptions worksheet. The users can change the values of these factors in the Assumptions worksheet and observe the impact of these changes on the stress test results.
- *Green* denotes numerical assumptions (parameters) of the stress test. Like the blue cells, the green cells are found only in the Assumptions worksheet. As with the blue cells, the users can change the values of these factors in the Assumptions worksheet and observe the impact of these changes on the stress test results.
- *No background.* Cells that have no background and generally normal black font, contain formulas linked to the yellow, green, and blue cells. If the values of the blue or green calls are changed (or if new input data are entered in the yellow cells), the results of the stress tests are recalculated automatically.
- *Yellow stripes* indicate consistency checks. These cells contain sums or other functions of the input data. Those would normally also come from the authorities as hard numbers, but are calculated in this file to avoid inconsistencies.
- *Green/white stripes* denote numerical assumptions imported from the Assumptions sheet. Under normal circumstances, it is expected that the user will leave these cells (each of

Table 1. Stress Tester 2.0: Description of the Worksheets

Worksheet	Description
Read Me	Explanation of the workbook.
Data	Six tables. Input data as compiled by the National Bank of Bankistan (NBB). The data were collected in March 2006 and generally relate to end-December 2005, unless noted otherwise. Table A1 contains basic balance sheet and income statement data. Table A2 contains other prudential indicators important for the stress tests. Tables A3 and A4 include key ratios based on the input data. Table A3 contains the FSIs, while Table A4 characterizes the structure of the banking sector. Tables A5 and A6 show how the FSIs can be combined into institution-by-institution rankings, using a supervisory early warning system calibrated by the NBB (see the Assumptions sheet). Table A5 provides the rankings; Table A6 converts them into probabilities of default.
Assumptions	One table. Table B puts together all the assumptions. This worksheet also contains several charts allowing the user to see how changes in the assumptions affect the results.
Credit Risk	Two tables. Table C1 summarizes the reported data on asset quality. Table C2 shows the credit risk stress test. It consists of four components: (1) a correction for underprovisioning of NPLs; (2) an aggregate NPL shock, (3) a sectoral shock, allowing different shocks to different sectors, and (4) a shock for credit concentration risk (large exposures).
Interest Risk	Two tables. Table D1 sorts assets and liabilities into three time-to-repricing buckets, using the input data provided by the NBB. Table D2 shows the corresponding interest rate stress test. The test itself consists of two components: (1) flow impact from a gap between interest sensitive assets and liabilities; (2) stock impact resulting from the repricing of bonds.
FX Risk	Two tables. Table E1 contains information on the foreign exchange exposure of the banks and the direct exchange rate risk shock. Table E2 shows a basic calculation of the indirect foreign exchange shock (using FX loans to approximate impact on credit quality).
Interbank	Three tables. Table F1 is a matrix of net interbank exposures. Table F2 uses the interbank exposure data to show "pure" interbank contagion, i.e. to illustrate what happens to the other banks when one bank fails to repay its obligations in the interbank market. Table F3 shows a "macro" contagion exercise, in which banks' failures to repay obligations in the interbank market are not assumed, but rather a result of the "macro" shocks modeled in the sheet "Scenarios."

Table 1. Stress Tester 2.0: Description of the Worksheets (continued)

Worksheet	**Description**
Liquidity	Two tables. The worksheet summarizes two liquidity tests, showing for each bank how many days it would be able to survive a liquidity drain without resorting to liquidity from outside (other banks or the central bank). Table *G1* models a liquidity drain that affects all banks in the system proportionally. Table *G2* is a model of "liquidity contagion," where the liquidity drain is faster in banks that are perceived to be similarly weak by depositors. This exercise also allows for testing the liquidity impact of government default.
Scenarios	Four tables. Table H1 summarizes the results of the combination of credit shocks, interest rate shocks, exchange rate shocks, and liquidity shocks from the respective worksheets. The table also compares the impact on profits and allows for an autonomous shock to profits. Table H2 shows the post-shock financial soundness ratios for the banking sector. Table H3 shows post-shock ratings. Table H4 shows the corresponding post-shock probabilities of default. The results, presented numerically in this worksheet, can be inspected visually in the "Assumptions" worksheet.

which contains a link to the “Assumptions” sheet) unchanged and will carry out the changes in the corresponding green cells in the “Assumptions” sheet. However, if users want to see the impact of changes in an assumption directly in the corresponding sheet (e.g., for credit risk assumptions in the “Credit Risk” worksheet), they can do it by changing the value of the green/white cell rather than going back to the “Assumptions” sheet. Of course, users need to be aware that that if they save these changes, it will result in overwriting the original links in the file and some of the links between the “Assumptions” sheet and the results in the “Scenarios” sheet may be broken. However, if they do not save those changes and afterwards return to the original template, they will be able to use the file again in its original form.

- *Blue/white stripes* denote numerical assumptions imported from the “Assumptions” sheet. Like the green/white stripe cells, these cells allow the user to change the values of assumed shocks directly in the individual worksheets without going back to the “Assumptions” sheet. However, these changes should be used with caution to avoid breaking the links in the file; changes in the shock sizes should primarily be done in the “Assumptions” sheet.

In addition to explanations in the “Read Me” sheet, many of the cells in the stress testing file contain comments explaining the calculations carried out in these cells.

All assumptions and shock parameters are in the "Assumptions" sheet. This is the sheet that a regular user would work with the most, changing the numerical assumptions (in green) and

shock sizes (in blue) and observing the results. Since a summary presentation of the stress test results is provided in the "Assumption" sheet in charts, the user can change the assumptions and directly see the impacts in a graphical form. If the user wants to examine the overall stress test results in a tabular form, the results are available in the "Scenarios" and various other sheets.

Expert users that have become familiar with the file are invited to suggest improvements in the file or to develop the file further themselves. Such developments can include new types of risks, making the modeling of the existing risks more realistic, or including more institutions in the system.[5] Not all the developments need to take place in the same file: users can think about some of the blue or green cells as interfaces between this tool (module) and other tools (modules), such as macroeconomic models that provide scenarios. Those can provide inputs that feed into this stress testing tool. The main advantage of Excel-based tools, such as this one, is the relative ease with which they can be adapted and extended. For longer-term usage, it may be useful to develop the file into a program, for example in MS Access. This may reduce the flexibility for regular users, but it may, among other things, allow development of the file from the current one-period snapshot to a multi-period framework.

B. Top Down or Bottom Up?

There are two main approaches to translating macroeconomic shocks and scenarios into financial sector variables: the "bottom-up" approach, where the impact is estimated using data on individual portfolios, and the "top-down" approach, where the impact is estimated using aggregated data.[6] Among central banks' financial stability reports (FSRs), reports by the Bank of England and Norges Bank can be used as examples of FSRs that rely more on top-down approaches to stress testing, while reports by the Austrian National Bank and Czech National Bank are examples of stress tests using more bottom-up approaches, even though in all these cases, the reports in fact combine elements from both approaches (see Appendix II for a more detailed overview of stress tests in FSRs).

The disadvantage of a top-down approach is that applying the tests only to aggregated data could overlook the concentration of exposures at the level of individual institutions and linkages among the institutions. This approach may therefore overlook the risk that failures in a few weak institutions can spread to the rest of the system. The bottom-up approach should be able to capture the concentration of risks and contagion, and should therefore

[5] Including more institutions requires adding columns (and some rows in the "Interbank" worksheet), copying the relevant links in other worksheets, and checking summation formulas for the peer groups and the system.

[6] For a longer discussion of this distinction, see e.g., World Bank and the International Monetary Fund (2005), Jones, Hilbers, and Slack (2004), and Čihák (2004a).

generally lead to more precise results, but it may be hampered by insufficient data and by calculation complexities. Having detailed information on exposures of individual banks to individual borrowers should in principle lead to more accurate results than using more aggregated data, but, especially for large and complex financial systems, it may lead to insurmountable computational problems. Most macroprudential stress tests therefore try to combine the advantages and minimize the disadvantages of the bottom-up and top-down approaches.

This document and the accompanying Excel file focus on the bottom-up implementation of stress tests in a relatively small, non-complex banking system. The spreadsheet illustrates why using institution-by-institution data is important: a relatively minor change in the distribution of risks among banks can result in substantial changes in the overall impacts. The spreadsheet also illustrates how the bottom-up approach can be complemented by a top-down approach. For example, a model estimated on aggregate data can be used to identify how a combination of shocks to macroeconomic variables can translate into an increase in nonperforming loans. Stress Tester 2.0 can then be used to calculate how this aggregate impact influences individual banks and the system as a whole.

The spreadsheet illustrates a “centralized” approach to bank-by-bank stress testing, where all the calculations are done in one center (e.g., at the central bank or a supervisory agency, or by an IMF expert). An alternative, more “decentralized” approach, often used in advanced country FSAPs (see Appendix III), is to involve banks themselves in carrying out the stress testing calculations. The advantage of the “decentralized” approach is that it can provide a richer, more detailed modeling, using a wider set of data, and leveraging on the expertise and calculation capacity of banks’ risk management.[7] The disadvantage is that the “decentralized” calculations may not sufficiently reflect contagion effects among banks, so just adding up the results for individual banks may not be a good proxy for the systemic impacts. Also, if the calculations are complex and done by many institutions, it may be a major challenge to ensure that all banks implement the assumed shocks or scenarios in a consistent fashion. In comparison, the “centralized” stress tests discussed in this paper are less refined (for reasons of computational complexity and data availability), but they (i) are more focused on linkages to macroeconomic factors; (ii) can better integrate credit and market risks; (iii) are implemented consistently across institutions; (iv) can better analyze correlation across institutions (inter-portfolio correlation); and (v) can analyze network effects (contagion). Therefore, even if the “decentralized” calculations are carried out, it is important to complement them with the “centralized” calculations along the lines discussed in this paper.

[7] Čihák and Heřmánek (2005) provide a more detailed discussion of such “decentralized” stress tests. Hoggarth, Logan, and Zicchino (2006) discuss how such stress tests were carried out in the case of the United Kingdom.

C. Presenting Stress Test Results—What Variables Can Be Stressed?

For a variable to be used to measure the impacts of the stress tests, it should have two key properties: (i) it should be possible to interpret the variable as a measure of financial soundness of the system in question, and (ii) it can be credibly linked to the risk factors. Of the various variables that have been used in the literature so far, each has advantages and disadvantages, which should be clear to the reader and user. Here is a list of the commonly used variables:

- *Capital.* Using capital as a measure of impact has a clear motivation. If a risk has a material impact on solvency, it has an impact on capital. Also, commercial banks' capital is part of the Other Items Net in monetary surveys, so expressing impacts in terms of capital could be used to directly link stress tests to other parts of the financial programming framework (even though it is only a small part of the possible feedback effects from the financial sector to the macroeconomy, as discussed in Section IX.C). The disadvantage of using impact in terms of capital is that it is just a number in B$ that needs to be compared to something else to give the reader an idea about the impact on soundness (e.g., dividing it by risk-weighted assets) and on the macroeconomic framework (e.g., dividing it by GDP). Nonetheless, it is a key measure, and the accompanying Excel file illustrates the presentation of stress testing impact in terms of capital.

- *Capitalization.* The advantage of capitalization measures (capital or equity to assets, or capital to risk-weighted assets) is that capital adequacy is a commonly recognized soundness indicator. Compared to capital, this measure is scaled, so it allows comparison among institutions of different size. For this reason, the accompanying Excel file uses capitalization as a key indicator of impact. The disadvantage of this measure is that a change in capitalization does not by itself indicate macroeconomic relevance of the calculated impacts. It therefore needs to be accompanied by other measures.

- *Capital injection needed* (for example as a percentage of GDP). This indicator provides a direct link to the macroeconomy. It provides an upper bound on the potential fiscal costs of bank failures associated with the assumed stressful scenario. The accompanying Excel file illustrates the use of capital injection.

- *Profits.* In a normal, non-stressful situation ("baseline scenario"), banks would typically create profits. When carrying out stress tests, it is important to bear in mind that we are evaluating impacts against such a baseline, as banks would normally use profits as the first line of defense before dipping into capital. Expressing shocks only in terms of capital may result in overestimating the actual impacts if banks were profitable in the baseline. The accompanying Excel file allows profits to be taken into account. More specifically, it indicates the "profit buffer" that banks would have available in the

baseline. The profit buffer is based on the average annual profits over the last 10 years, but the file also allows the user to run a separate "test" for the impact of an autonomous shock affecting profits (one can think about autonomous shocks to net interest income, e.g., related to an increase in competition from abroad; alternatively, this can be viewed as a measure of the risks not reflected in the other shocks specified in the stress testing scenario). However, to reflect the views of some observers that it is better to be prudent and disregard profits, it shows them as a separate item rather than directly deducting the impacts from profits.

- *Profitability* (return on equity, assets, or risk-weighted assets). Compared to profits, these measures are scaled by bank size, thus allowing comparison among banks of different size. The accompanying Excel file shows the profit buffers as ratios to the risk-weighted assets, to make them comparable with the impact shown in terms of risk-weighted assets.

- *Net interest income and other components of profits.* Sometimes it can be useful to stress test separately individual components of profits. For example, net interest income is likely to have a more direct relationship to interest rates, and may therefore be more amenable to econometric analysis. However, such an approach provides only a partial picture of the economic value of a bank and its resilience to adverse events.

- *Z-scores*. The z-score has become a popular measure of bank soundness (see, e.g., Boyd and Runkle, 1993; or Hesse and Čihák, 2007). Its popularity stems from the fact that it is directly related to the probability of a bank's insolvency, i.e. the probability that the value of its assets becomes lower than the value of the debt. The z-score can be summarized as $z \equiv (k+\mu)/\sigma$, where k is equity capital as percent of assets, μ is average after-tax return as percent on assets, and σ is standard deviation of the after-tax return on assets, as a proxy for return volatility. The z-score measures the number of standard deviations a return realization has to fall in order to deplete equity, under the assumption of normality of banks' returns. A higher z-score corresponds to a lower upper bound of insolvency risk—a higher z-score therefore implies a lower probability of insolvency risk. The accompanying Excel file illustrates this presentation.[8]

- *Loan losses.* Stress tests presented by staff from Norges Bank (e.g., Evjen and others, 2005) and Bank of England (e.g., Bunn, Cunningham, and Drehmann, 2005) are just two examples of presenting stress testing results in terms of loan losses. While this approach has its advantages (in particular, it is easier to implement in top-down

---

[8] The file even shows the z-scores for the peer groups of banks, using a similar methodology as used by some authors to translate bank-by-bank distance to default measures to a "portfolio distance to default" (e.g., IMF, 2005a). However, these calculations need to be treated with a degree of caution, given that they overlook issues related to contagion (see, e.g., Čihák, 2007).

calculations than measures calculating losses to capital), its drawback is that it does not take into account banks' buffers (profits and capital) against those losses. It may underestimate the overall impact if losses are concentrated in weak institutions.

- *Liquidity indicators.* For liquidity stress tests, the impacts have to be measured differently than for solvency tests, namely in terms of liquidity indicators. The accompanying Excel file illustrates this presentation.

- *Ratings and probabilities of default (PDs).* Ratings and PDs provide a useful way of combining solvency and liquidity risks. By definition, ratings try to combine various solvency and liquidity risks into a single measure. We can use the system designed for ratings and see how changes in the various variables translate into changes in ratings. If we have a model linking ratings and probabilities of default, we can also calculate how a stressful scenario influences PDs. The accompanying Excel file illustrates this presentation.

This list is far from complete. It is possible to calculate and present stress test impacts in terms of other variables that capture soundness of financial institutions and can be credibly linked to the development of risk factors. For example, instead of the accounting-based data discussed above, it is possible to present the impact of a stressful scenario in terms of market-based indicators of financial sector soundness, such as relative prices of securities issued by financial securities, the distance to default for banks' stocks,[9] or credit default swap premia (for a review, see e.g., Čihák, 2007). One of the advantages of the market-based indicators is that they are usually available on a much more frequent basis than accounting data. However, one of their major disadvantages is the absence in many countries of sufficiently deep markets from which such indicators can be derived (e.g., bank stocks are not traded or the market for such stocks is illiquid). There have been some attempts to link market-based indicators, such as distance to default, to macroeconomic variables (see, e.g., IMF, 2005a), but work on using such relationships for stress tests has been limited so far.

D. How Are the Results Presented in Stress Tester 2.0?

Each stress test conducted in this exercise aims to address two main questions: (1) Which banks could withstand the assumed shocks and which ones would fail? (2) What are the associated potential costs for the government given the failure of banks in times of stress?

[9] Distance to default is in effect an implementation of the z-score, discussed above, for banks with stocks listed in liquid equity markets. Distance to default uses stock price data to estimate the volatility in the economic capital of the bank (see, e.g., Danmark Nationalbank, 2004).

A commonly used approach to assessing question (1), as mentioned in the previous section, is to look at banks' capital adequacy ratio (CAR). According to the original Basel Agreement, a bank has to hold a minimum CAR, defined as total regulatory capital to risk-weighted assets (RWA), of 8 percent. In our example, given that Bankistan as an emerging market country faces more risks than an industrial country, it is appropriate to require from banks a higher CAR. Based on these considerations, Bankistan's supervisors use a minimum CAR of 10 percent. Whenever the CAR of a bank in Bankistan falls below 10 percent, its owners are obliged to inject capital in order to stay in business.[10] If they fail to do so, the bank will be closed and its banking license withdrawn. If the CAR is below 0, then the bank has negative capital and is insolvent (and, of course, needs to be closed or capital has to be injected).[11]

Stress Tester 2.0 allows profits to be taken into account, which—as explained earlier—is important since shocks always take place over time, and therefore need to be compared with a "baseline scenario". The "Scenarios" sheet indicates the profit buffer that banks would have available in the baseline scenario. We refer to annual profits, which is consistent with the fact that we evaluate the shocks in a horizon of one year (see the interest rate shock). The value of the profit buffer is based on the average annual profits over the last 10 years, but the file also allows the user to run a separate "test" for the impact of an autonomous shock affecting profits or net interest income. Take, for instance, autonomous shocks to net interest income, e.g., related to an increase in competition from abroad; alternatively, this can be viewed as a measure of the risks not reflected in the other shocks specified in the stress testing scenario. However, to reflect the views of some papers that it is more prudent to disregard profits and measure shocks directly against capital (e.g., Blaschke and others, 2001), the file shows the profit buffers as a separate item rather than directly deducting the impacts from profits. For some banks, the profit "buffer" is non-existent or negative (they have been creating losses).

An assessment of question (2) requires consideration of the following question: If bank owners fail to inject new capital, how much capital would the government need to inject in order to bring the CAR up to 10 percent again? For state-owned banks, it is obvious that the government would have to inject capital to keep banks operating. For private-owned banks, asking this question assumes that the government has an implicit or explicit guarantee for the banking sector, which may or may not be the case. If it is, an answer to this question can be found from the following accounting relationship:

---

[10] The 10 percent minimum capital adequacy requirement is just an **assumption** in Stress Tester 2.0, contained in the B71 cell of the "Assumptions" worksheet. The reader can change it, for example, to 8 percent, and see in the relevant chart in the "Assumptions" worksheet how the capital injection (in percent of GDP) declines.

[11] Total capital may in general differ from regulatory capital; in this workbook, we use for simplicity the same numbers, but the file is set up in a way that allows for differences between equity and regulatory capital.

$$\frac{C+I}{RWA+qI}=\rho$$

where C is the bank's existing total regulatory capital, RWA are its existing risk-weighted assets, I is the capital injection, q is the percentage of the capital injection that is immediately used to increase risk-weighted assets, and ρ is the regulatory minimum CAR (ρ =10 percent in the case of Bankistan).

From the above equation, we can express the necessary capital injection as:

$$\begin{aligned} I &= \frac{\rho\ RWA - C}{1-q\rho} \quad if\ C < \rho\ RWA; \\ &= 0 \quad otherwise \end{aligned}.$$

If q=0, i.e. the capital injection is not used for an increase in RWA (at least immediately), and **if** we substitute for ρ the 10 percent value used in Bankistan, we can calculate the capital injection as I = 0.1*RWA – C. The values of the parameters ρ and q are **assumed**. The stress test file enables us to change the values of the two parameters in the appropriate green cells (B71 and B72 in the worksheet "Assumptions"). We can see that if ρ is lower than 10 percent, the necessary capital injection is lower, and vice versa. If RWA increase as a result of the increase in capital (i.e. if q>0), the necessary capital injection is higher (but the impact of changes in q is generally rather small).

III. Understanding And Analyzing the Input Data

The "Data" sheet summarizes the input data (Tables A1 and A2), as reported by the NBB. It also shows key ratios (Tables A3 and A4), and illustrates how these ratios can be used in an off-site supervisory assessment system (Tables A5 and A6).

A. Coverage of Stress Tests

In the accompanying Excel example, the stress tests cover all 12 commercial banks in the country. An overview of FSAP stress tests suggests that only a minority of FSAPs have covered all banks in a country. Most FSAPs have covered only a subsample of large banks, accounting for a substantial majority (generally 70–80 percent) of the banking system's total assets. Including all banks rather than a subsample has the obvious advantage of being more comprehensive. For this reason, this approach is also often favored by supervisors, who are expected to supervise all institutions, not only the larger ones. However, for someone interested primarily in macroprudential issues (e.g., a central bank not involved in microprudential supervision), it may be sufficient to include only the systemically important institutions (that is why some authors, such as Jones, Hilbers, and Slack (2004) refer to this type of stress testing as "system-oriented stress testing" rather than "system-wide stress

testing"). Excluding the other institutions may be practical for reasons of computational complexity.[12]

The Stress Tester 2.0 file illustrates a basic peer group analysis. To do that, the 12 banks in Bankistan are grouped into 3 peer groups according to their ownership: state-owned banks (SB1 to SB3), domestic privately-owned banks (DB1 to DB5), and foreign-owned banks (FB1 to FB4). This is just one possible grouping. Depending on the analytical purpose, the banks in Bankistan can be grouped into other groups, for example by their size (e.g., large, medium, and small) or financial performance (e.g., strong and weak).

In the Bankistan example, foreign banks are present only through locally-incorporated subsidiaries. In practice, foreign banks may be present in the market directly, through their branches. The practical difference from a stress testing perspective is that branches typically do not have their own capital against which the impact of shocks could be shown (and comparing the impacts with the parent institution's capital would be misleading, because it would overlook risks faced by the same institution in other countries). However, if there are separately available data on assets, liabilities, incomes, expenses, and other input data indicated in the "Data" worksheet, most of the standard tests can be performed on the branches of foreign banks as well. The main practical difference then is that the impacts have to be expressed in terms of a different variable than capital or capital adequacy (e.g., in terms of profits).

Mirroring the practice of FSAPs and central banks' financial stability reports (FSRs), this document focuses on stress tests for banks and banking systems. In most countries, banks tend to dominate the financial system and are key to assessing systemic risk. Some FSAPs and FSRs contained explicit stress tests of insurance companies and pension funds (see Appendices II and III, respectively).
Stress testing pension funds and insurance companies themselves can be a rather complex task. As an illustration, Box 2 provides an overview of the various risks, which can be subject to stress testing in insurance companies. Some of these risks, such as market risk, liquidity risk, or credit risk, can be modeled similarly to the stress testing of banks. Modeling of other risks,

---

[12] A practical complication with the system-oriented stress tests is the fact that the systemic relevance of an institution is established only after, not before the stress tests. As a practical shortcut, many FSAPs and other authors use a measure of size (e.g., total assets) as a first approximation for systemic importance. It is a good proxy in most cases, but in some it can miss institutions that are small but have large potential for impact on other institutions, e.g., through exposures in the interbank market. In Section VII, we provide a measure of systemic importance that takes the contagion effects into account.

Box 2. Stress Tests for Insurance Companies

The risks that can be addressed by insurance sector stress testing include the following:[1/]

- *Underwriting risk*. This includes risks associated with rapid growth or decline in the volume of the underwriting portfolio, uncertainty of the claims experience, length of tail of the claims development, dependence on intermediaries, possibility of reinsurance rates increasing substantially, effects of a high level of uncertainty in pricing in new or emerging underwriting markets, geographical mix of the portfolio, and tolerance for variations in expenses.
- *Catastrophe risk.* This risk reflects the ability of the insurer to withstand catastrophic events, increases in unexpected exposures, latent claims or aggregation of claims, or the possible exhaustion of reinsurance arrangements, and the appropriateness of the catastrophe models and underlying assumptions used.
- *Deterioration of technical provisions.* This includes the adequacy and uncertainty of the technical claims provisions, the adequacy of other underwriting provisions, the frequency and size of large claims, possible outcomes relating to any disputed claims, particularly where the outcome is subject to legal proceedings, the effects of inflation, the effects of increasing longevity on pension products, the guarantees and options in policy terms, the risks of early policy termination which can be linked to variations in interest rates, social changes resulting in an increase in the propensity to claim or to sue, other social, economic, legislative and technological changes.
- *Market risk.* This reflects adverse movement in the value of an insurer's assets and liabilities affected by market movement. Modeling this risk is not much different from banking.
- *Credit risk.* This involves the failure of counterparties (debtors, brokers, policyholders, reinsurers, guarantors) to perform on obligations. Also in this case, modeling the risk is similar to banking.
- *Liquidity risk.* This relates to the possibility that an insurer will be unable to realize assets to fund its obligations as and when they fall due. Again, modeling the risk is similar to banking.
- *Other risks.* These include operational risk, group risk, and systemic risk (e.g., impact of failures/downgrades in other insurers or banks).

1/ This list is based broadly on IAIS (2003).

especially some of those stemming from the liabilities side (e.g., catastrophe risk), would be beyond the scope of this paper.

The impact of failures in nonbank financial institutions can be assessed as part of the credit risk. The accompanying Excel file allows for two ways of incorporating such an analysis. First, it can be incorporated as part of the sectoral credit risk, with nonbank financial institutions as one of the sectors. We can then run a basic stress test for what would happen if

a certain percentage of loans to the nonbank financial sector became nonperforming. Second, it can be incorporated as part of the large exposures tests. If we have data on the largest exposures of banks to nonbank financial institutions, we can run a test on what would happen to banks' solvency should their largest counterparties in the nonbank financial sector fail.

B. Balance Sheets, Income Statements, and Other Input Data

Data availability is a key determinant of the quality of a stress test. The purpose of Tables A1 and A2 in the "Data" worksheet is to illustrate what data are typically needed for carrying out a set of stress tests. These are not the minimum requirements—it is possible to do very rudimentary stress tests with even less data, as discussed in Box 3. However, these are the types of data that one usually looks for when doing a basic stress test.

Throughout the sheet, as well as the file, the first data column shows aggregated data for the whole banking system, the next three columns show the aggregated data for the three peer groups of banks (state banks, private domestic banks, and foreign banks), and the remaining twelve columns show the data for the individual banks.

The input data presented in the top part of the "Data" worksheet (Table A1) form a set of balance sheets and income statements. To keep the exercise straightforward and transparent, the aggregated data and the peer group data can be calculated as sums of the bank-by-bank data. This means that we disregard interbank exposures for the time being. This assumption is relaxed later on, in the analysis of interbank contagion risk (Section VII).
Table A2 in the "Data" worksheet lists—again for the system in aggregate, for the peer groups, and for individual banks—other key input data that are used for the stress test calculations. Such data are on the following: (i) regulatory capital and risk-weighted assets (to be able to express the stress test impacts in terms of capital adequacy), (ii) asset quality and structure of lending by sector and by size of borrower, (iii) provisioning and collateral (for the credit risk calculation), (iv) structure of assets, liabilities, and off-balance sheet items by time to repricing, (v) the structure of the bond portfolio (for the interest rate risk calculation), (vi) net open positions in foreign exchange and lending in foreign currency (for the foreign exchange solvency risk calculation), (vii) average profits and standard deviation of profits over time (to have a measure of "baseline" profitability), (viii) liquidity structure of assets and liabilities (for liquidity risk calculation), and (ix) bank-to-bank uncollateralized exposures, presented in matrix form (for the interbank solvency contagion risk).

Most of the data in Table A2 are usually available to bank supervisors through standard regulatory returns. However, several issues need to be mentioned:

- Stress tests analyze the economic position *(net worth) of banks*. This should in principle be aligned with the reported data on capital, but in practice there may be important differences between the calculated economic net worth of a bank, and the reported

regulatory data on capital. This may be the case, for example, when some assets are overvalued in banks' balance sheets, or when regulators accept as capital some liabilities that in fact are not capital (e.g., some long-term loans). The person carrying out stress tests should first try to adjust the input data for such biases. In the file, we show one example of such adjustments, namely when banks underprovision their nonperforming loans. Another example (not shown in the file) might be when banks do not mark to market some bonds that they are holding in their portfolios.

- The input data should reflect not only assets and liabilities, but also *off-balance sheet* positions. For example, the net open positions in foreign currency should reflect the delta equivalents of foreign exchange options.

- Data on *bank-to-bank exposures* may be difficult to collect in many countries. In those cases, approximate calculations using less data (e.g., only data on each bank's exposure to the rest of the system as a whole) can be used to at least broadly assess the associated risks, even though such methods may result in overlooking some exposures in the system.

- For interest rate risk calculations, it is important to have data on *time to repricing*. For example, from an interest rate risk perspective, a 20 year mortgage loan with an interest rate that can change every 6 months, should be treated the same as a 6 month fixed rate loan, not as a 20 year loan. However, getting data on time to repricing may be difficult to obtain in some cases. In many countries, banks report instead a breakdown of assets by maturity or by residual maturity. While data on maturity or residual maturity are important for analyzing liquidity, using them as proxy for time to repricing may lead to misleading results (typically overstating the interest rate risk).

Additional data may be needed for stress testing for other risks not explicitly covered in this file. For example, to carry out stress tests for equity price risk and commodity price risk, data on net open positions in equities and in commodities would be needed (the mechanics of the test would be similar to the direct foreign exchange solvency test). Also, breakdowns of assets and liabilities by residual maturity/time to repricing **and** currency would be needed to perform stress tests separately for foreign currency.

Box 3. How To Do Stress Tests When NPLs or Other Input Data Are Unavailable?

What if some of the input data in the Stress Tester 2.0 exercise, such as those on nonperforming loans, net open positions, and time to repricing are not available to the analyst? This may be for example because of weak supervisory reporting systems or because of legal restrictions on data sharing.

If some of the data are not available, a more rudimentary version of the stress tests can still be performed with just the basic financial statements, provided that they are available for a sufficient number of periods. Those can be based on the observed (or assumed) relationships between the risk factors, the various items of the income statement, and the balance sheet. For example, even when no data are available on the repricing buckets of assets and liabilities and on the bond portfolios in banks, a very rudimentary stress test for interest rate risk could be based on the net interest income on banks' income statements. In particular, past data on net interest income of individual banks over time can be regressed on interest rates and other potential variables to estimate how banks' net interest income responded to changes in interest rates, and the estimated slope coefficient(s) can be used to translate a change in interest rates into the impact in terms of profits (and potentially capital). Similarly, provisions for loan losses from the income statement can be regressed on the risk factors and other explanatory variables to analyze the impact on banks' profitability. If sufficiently long time series are not available to carry out such regressions in a country, the slope coefficients can be calibrated based on expert information or experience from other countries.

Even if the individual items of the financial statements are not available, "reduced-form" stress tests can still be carried out if there are reliable time series data. For example, one needs just capital, asset, and return data on individual banks over time to calculate the z-score, as a proxy for individual bank soundness, discussed in section III.C. The z-scores for individual banks can be regressed on a range of macroeconomic variables (e.g., real GDP growth rate, interest rate, and exchange rate) and bank-level variables (e.g., asset size or loan to asset ratio). The slope coefficients from this regression can then be used to map a macroeconomic scenario into the z-scores, to approximate the impact of macroeconomic stress on individual bank soundness. (A similar approach can also be used with distance to default data or other market-based indicators of individual institution soundness). The main challenge in this type of approach is how to aggregate the bank-by-bank soundness data into a system-wide indicator (an issue that is discussed in more detail in Čihák, 2007).

C. Indicators of Financial Sector Soundness and Structure

Table A3 contains a set of core FSIs and other important ratios characterizing Bankistan's banking sector and its components. These ratios can be used to provide a summary picture of

the soundness of the financial sector and its components (peer groups and individual banks) in Bankistan. For more on the definitions and compilation of FSIs, see IMF (2004).

Table A3 also includes, at the bottom, individual banks' z-scores. The z-score is defined as $z \equiv (k+\mu)/\sigma$, where k is equity capital as a percent of assets, μ is average after-tax return as percent on assets, and σ is standard deviation of the after-tax return on assets, as a proxy for return volatility. As mentioned earlier, z-scores have become popular as measures of bank soundness, because they are directly linked to the probability of a bank's insolvency. The table also shows the z-scores for the peer groups of banks, using a similar methodology as used by some authors to translate bank-by-bank distance to default measures to a "portfolio distance to default" (e.g., IMF, 2005a). These calculations need to be treated with caution, because they overlook issues related to contagion (see, e.g., Čihák, 2007).

Table A4 characterizes the structure of the banking sector in Bankistan, showing the shares of the peer groups and the individual banks in total assets, loans, deposits, and capital. It shows that the share of foreign-owned banks is relatively high—about 55 percent in terms of assets, and 83 percent in terms of capital (reflecting their higher capitalization). It also shows the total assets as a ratio to gross domestic product (GDP), to indicate the size of the banking system relative to the Bankistan economy. The GDP figure is used again later, to put into macroeconomic perspective the capital injection needed to get all banks to comply with the minimum capital adequacy requirement.

D. Ratings and Probabilities of Default

Tables A5 and A6 show how banking sector ratios can be combined into institution-by-institution rankings, using a supervisory early warning system.[13] Such systems (e.g., Sahajwala and Van den Berg, 2000) are very common in supervisory agencies, and they are typically used for assessing soundness of banks in "baseline" conditions; we will show in this exercise that they can be also used for assessing soundness in stressful conditions. One caveat needs to be borne in mind, however, namely that these early warning systems typically treat each bank separately and do not look at contagion among banks—an issue that will be analyzed as part of the stress test exercise presented here.

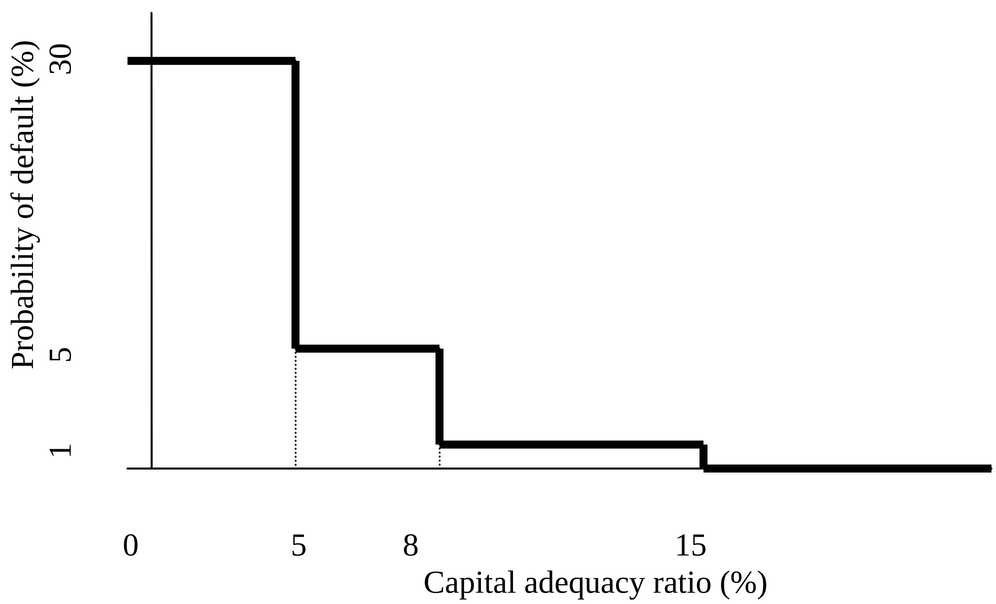
Table A5 provides the rankings, based on the off-site supervisory ranking system of the NBB, characterized in rows 3–22 of the "Assumptions" worksheet. The system has three thresholds (columns B, C, and D in the "Assumptions" worksheet) for each indicator, determining a numerical ranking to each of the indicators (with 1 indicating the best ranking,

[13] The early warning system does not have to comprise only ratios. Inclusion of variables in the early warning system should reflect their power to identify weak banks. One variable that in some cases has good discriminatory power is deposit rates: high deposit rates in a bank can indicate that it has difficulty retaining depositors, a potential sign of problems (Kraft and Galac, 2006).

and 4 the worst ranking). The rankings for the individual variables are weighted (using weights established by the NBB and provided in column E of the "Assumptions" worksheet) to derive an overall ranking for a bank. Similar off-site ranking systems are used by supervisors to identify banks that deserve increased attention (Sahajwala and Van den Berg, 2000).

Underlying each ranking system is a more or less explicit link to probability of default (or probability of technical insolvency, i.e. probability that the capital adequacy ratio declines below the regulatory minimum). Table A6 illustrates this by converting the rankings into probabilities of default, using a "step function" illustrated in Figure 2, and described in the case of Bankistan rows 6 and 22 of the "Assumptions" worksheet: according to NBB's estimates, a bank with a rating of 1 has a 0.1 percent probability of default in a given year; a bank rated 2 fails with a 1 percent probability; a bank rated 3 has a 5 percent probability of default, and a bank rated 4 has a 30 percent probability of default in the coming year.

Figure 2. 'Step Function' (Example)



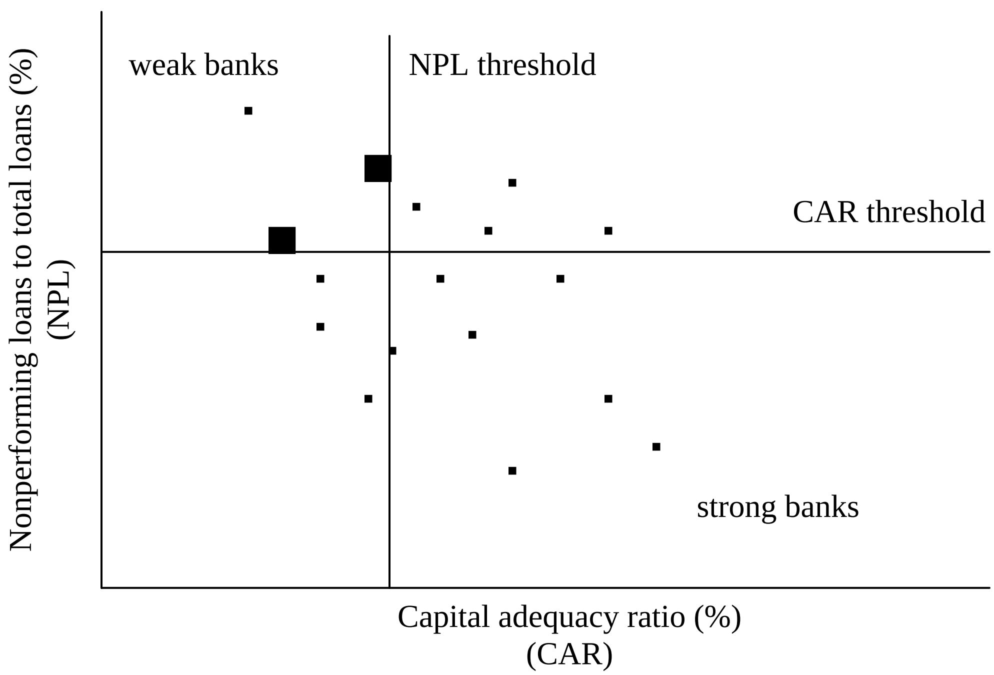
Source: the author, based on the default settings in the "Assumptions" sheet

How are the parameters of such step functions derived? In some cases, they are based on expert estimates. In other cases, central banks or supervisory agencies attempt to "back-test" such systems to see if they actually identify institutions that fail (or institutions that need

interventions). When the step function is re-estimated, the new parameters should be entered in row 22.

Figure 3 provides an illustration of such "back-testing" in the case of two variables (capital adequacy and gross nonperforming loans to total loans) and one threshold per variable. The dots represent observations of banks, and the two bigger boxes indicate two banks that have actually failed. The supervisory ranking system attempts to single out the failed banks (i.e., the bigger boxes), minimizing the signal-to-noise ratio for the estimate. If we want to capture all failed banks in Figure 3 (i.e., eliminate Type I errors), the early warning system characterized by the CAR threshold and the NPL threshold shown in the figure allows to decrease the percentage of banks misclassified as failures (i.e., Type II errors) from 88 percent (15/17, if we do not have any prior information) to 33 percent (1/3, for the "northwest" sub-set identified by the two thresholds). That is a major improvement in forecast precision.

Figure 3. Back-Testing a Supervisory Early Warning System (Example)



IV. Credit Risk

Lending is the core of the traditional banking business. In most banking systems, credit risk is the key type of risk. At the same time, it is the type of risk where existing models are most in need of strengthening.

There are three basic groups of approaches to modeling credit risk as part of stress tests. First, there are mechanical approaches (typically used if there are insufficient data or if shocks are different from past ones). Second, there are approaches based on loan performance data (e.g., probabilities of default, losses given default, nonperforming loans, and provisions) and regressions (e.g., single equation, structural, and vector auto regression). Third, there are approaches based on corporate sector data (e.g., leverage or interest coverage) and possibly on household sector data (even though such data are typically much more difficult to collect than corporate sector data).

The exposition in this section and in the accompanying file starts with the basic mechanical approaches. We then discuss how these can be extended into more realistic approaches. In particular, Box 4 has a discussion of the links between credit risk and macroeconomic risk.

Worksheet “Credit Risk” contains calculations relating to the credit risk of banks, i.e. the risk that banks’ borrowers will default on their contractual obligations. Table C1 in the worksheet “Credit Risk” summarizes the reported data for asset quality. Table C2 is used for the credit risk stress tests. It includes four different types of credit shocks, labeled 1, 2, 3, and 4.

A. Credit Shock 1 (“Adjustment for Underprovisioning”)

The purpose of the first part of the credit risk calculation is to make the point that stress testing should focus on the underlying economic value (net worth) of the bank. The economic value may in general differ from the bank’s reported regulatory capital. For example, as part of reported capital, some banks may include items that in fact are not capital and should rather be treated as something else (e.g., long-term loan). Or they can overstate some assets, resulting in overstated capital. Because all the stress testing calculations relate to the economic value of the bank, the stress testing analyst needs to first adjust the reported data to get a better picture of the starting (“baseline”) economic situation of the bank. Credit shock 1 shows an example of such an adjustment.[14]

[14] In a strict sense, therefore, this is just a starting point adjustment, not a part of stress tests themselves. In a broader sense, however, it is a “stress test” that shows how the reported data would change if the reporting system changed to one that reflected more closely the economic value of a bank. Of course, if the reporting system already reflects the economic value, there is no adjustment, and one can proceed directly to Credit Shock 2 and beyond.

Box 4. Linking Credit Risk and Macroeconomic Models

A number of papers have attempted to link credit risk to macroeconomic variables using econometric models. For example, Pesola (2005) presents an econometric study of macroeconomic determinants of credit risk and other sources of banking fragility and distress in the Nordic countries, Belgium, Germany, Greece, Spain and the UK from the early 1980s to 2002. An even broader cross-country analysis is presented in IMF (2003). For Austria, Boss (2002) and Boss and others (2004) provide estimates of the relationship between macroeconomic variables and credit risk. For Finland, Virolainen (2004) develops a macroeconomic credit risk model, estimating the probability of default in various industries as a function of a range of macroeconomic variables. For Norway, the Norges Bank has single equation models for household debt and house prices, and a model of corporate bankruptcies based on annual accounts for all Norwegian enterprises (Eklund, Larsen, and Berhardsen, 2003). For Hong Kong SAR, two studies are available on the topic, one being a single equation aggregate estimate (Peng and others, 2003), and one being a panel using bank-by-bank data (Gerlach, Peng, and Shu, 2004). For the Czech Republic, Babouček and Jančar (2005) estimate a vector autoregression model with nonperforming loans and a set of macroeconomic variables.

Similar models are also common in Financial Sector Assessment Program (FSAP) missions. For example, the technical note from the Spain FSAP includes an estimate of a regression explaining nonperforming loans on an aggregate level with financial sector indicators and a set of macroeconomic indicators (IMF, 2006). Appendix III has more details on stress tests in FSAPs.

Several issues need to be considered when interpreting the macroeconomic models of credit risk. In particular, the literature is dominated by linear statistical models. The linear approximation may be reasonable when shocks are small, but non-linearities are likely to be important for large shocks: doubling the size of the shock may more than double its impact. Indeed, micro-level credit risk models often find a non-linear relationship between the scale of shocks and the likelihood of default; for macroeconomic shocks, Drehmann (2005) also reports a nonlinear link to credit risk. Moreover, the models are subject to the Lucas critique (Lucas, 1976), since their parameters or functional forms may become unstable, especially if exposed to a major stress. As an extreme example, when considering a scenario that involves de-pegging in a country with a currency board regime, models estimated on past data are likely to say very little about the impact of the exchange rate change on credit risk. In such a situation, other approaches, such as calibration that uses parameters based on experience from other countries, may be more appropriate.

In Credit Shock 1, we assess what would happen if banks corrected their currently insufficient provisioning to fully meet the existing provisioning requirements. As mentioned earlier, a BCP assessment suggested that the current loan classification and provisioning standards in Bankistan are broadly adequate, but not well enforced. The regulations prescribe the following loan provisioning rules: 1 percent general provision for pass loans, 3 percent

general provision for special mention loans, 20 percent specific provision for sub-standard loans, 50 percent specific provision for doubtful loans, and 100 percent for loss loans). Reflecting the fact that it is very difficult to foreclose collateral in Bankistan, the test **assumes** that the actual value of the collateral is only 25 percent of the reported values (i.e., we assume a 75 percent "haircut" on the value of the collateral).

B. Credit Shock 2 ("Increase in NPLs")

If Credit Shock 1 is a "starting point adjustment," then Credit shock 2 can be seen as the first "real" stress test. It models a general decline in asset quality, affecting all banks proportionately. It is **assumed** that nonperforming loans (NPLs) increase by a certain percentage, the default values being 25 percent of the existing stock of NPLs. This means that a bank would have to undertake additional provisioning by 25 percent for each of the three groups that constitute NPL (substandard, doubtful, and loss loans). The increased provisioning requirements will reduce the value of the RWA as well as the capital. As regards the impact on the RWA, a common **assumption** is that the full increase in NPL is subtracted from RWA. However, the impact on RWA may be smaller if the impacted assets have a weight of less than one in the RWA. Typically, precise information of the distribution of NPL across risk-categories is not available. Nonetheless, a green cell (B40) in the "Assumptions" worksheet allows the user to change the **assumed** weight from 100 percent to a smaller number (say, 80 percent) and observe the impact of this assumption on the results.

The initial **assumption** in credit shock 2 is that the increase in NPLs in individual banks is proportional to the existing NPLs in these banks. In other words, banks that had more NPLs in the past, are assumed to have more new NPLs as a result of the shock. This is the most straightforward calculation, but there are alternative approaches. For example, the new NPLs can be proportional to the overall stock of loans, or to the stock of new performing loans. The green cells in this part of the worksheet allow the relaxation of this **assumption** and choose the weight of the existing NPLs and performing loans in determining the bank-by-bank increases in NPLs.

Whether to use existing NPLs or existing performing loans as a basis for assessing future credit risk is an open question. It is an empirical question that—if there are sufficient empirical data—can and should be decided empirically, by testing for the factors explaining bank-by-bank changes in NPLs (see Box 4 for a discussion of this issue). In the absence of reliable or sufficiently detailed empirical data, however, it is often necessary to resort to simplifying assumptions. As a rule of thumb, using the existing NPLs is correct if the existing bad loans are a good proxy for the quality of a bank's risk management and therefore of the risk faced by the bank going forward. Conversely, using the existing performing loans can be justified by the fact that performing loans are those loans that may "go bad," i.e. be shifted into the NPL category, and therefore indicate the potential for credit risk. Using performing loans as a basis may be warranted if there has been a structural change in the

economy (and therefore the past NPL ratios are of limited guidance in assessing future credit risks). For example, in a number of emerging markets in Central and Eastern Europe in the early 2000s, there has been a marked shift from corporate lending to household lending, with very different qualities and parameters. Using past NPLs as a basis for the bank-by-bank increases in such a situation might lead to misleading results.

The importance of the bank-by-bank distribution of NPL increases can be illustrated in the "Credit Risk" sheet. For example, in the default settings, NPLs increase by 25 percent in each bank (i.e., cell B46 in the "Credit Risk" sheet has a value of 25, and cells B48 and B49 have the values of 1 and 0, respectively).[15] The total volume of additional NPLs in the system (cell B50) is B$ 2,206 million and the state-owned banks' post-shock capital B$ -437 million (cell C53). If we instead make the increase in NPLs proportional to the existing performing loans (i.e., we make cells B48 and B49 equal 0 and 1, respectively) and make the overall volume of additional NPLs the same (which we can achieve by using the Tools/Goal Seek function, setting the value of B50 back to B$ 2,206 million by changing the values of B46), the increase in NPLs has to be about 4.3 percent of performing loans, and the state-owned banks' post-shock capital is B$ -296 billion (cell C53), i.e. smaller than if the increase in NPLs is proportional to the stock of NPLs. So, if we are concerned about the fiscal costs of the state-owned banks, the distribution of credit risks and buffers is important.

C. Credit Shock 3 ("Sectoral Shocks")

As an illustration of how bank-by-bank credit risk can be modeled in a more realistic fashion, the stress testing file includes also sectoral shocks (see the bottom part of Table C2).

This exercise allows the user to select different shocks to economic sectors and observe how each bank would be impacted, depending on the relative sizes of the banks' credit exposures to these sectors. The case shown in the file as a starting example models a "terrorist attack" scenario, increasing credit risk in the tourism and trade sectors. The calibration of the sectoral shocks can be based on a historical scenario (e.g., a concrete example of a terrorist attack in Bankistan or in a neighboring country) or on empirical models explaining, based on past data, default rates in different sectors as a function of macroeconomic and other explanatory variables. The B43–B49 cells in the "Assumptions" worksheet can then be seen as an interface between the econometric model (or historical scenario) and the balance sheet implementation calculated in Stress Tester 2.0.

The definitions of sectors may be adapted depending on the country and the analyzed topic. For example, the "other" sector may be broken down into a number of sub-sectors. Alternatively, instead of using the main economic activity of the counterparty as the defining

[15] These values originate in cells B35, B37, and B38 of the "Assumptions" worksheet.

feature, sectors can be defined by the nature of the counterparty. For example, the "sectors" can be households (residents/nonresidents), nonfinancial enterprises (residents/nonresidents), nonbank financial institutions, and government (domestic/foreign).

The increase in NPLs is **assumed** to be proportional to the particular bank's credit exposure to a sector, approximated by the bank's total loans to that sector. Along the lines of the discussion in the previous sub-section, it is possible to envisage that the increases reflect the bank's existing NPLs or total loans to each sector, and relax this assumption accordingly.

D. Credit Shock 4 ("Concentration Risk")

Stress Tester 2.0 incorporates a block at the end of Table C2 that allows testing for the failure of the largest counterparties of individual banks. The user can change the **assumed** number of failures per institution, and the **assumed** provisioning rate for those failures. If supervisors have data on banks' exposures to nonbank financial institutions (NBFIs), this type of test can also be used to model the credit impact on banks of failures of the largest NBFIs.

V. Interest Rate Risk

The interest rate shock in the worksheet "Interest Risk" tests for direct interest rate risk. Direct interest rate risk is the risk incurred by a financial institution when the interest rate sensitivities of its assets and liabilities are mismatched. In addition, the financial institution is also exposed to indirect interest rate risk, resulting from the impact of interest rate changes on borrowers' creditworthiness and ability to repay. The indirect interest rate risk is a part of credit risk. We discuss how to account for the interest rate-related credit risk in the section devoted to designing scenarios involving several shocks (Section IX).

A. Direct Interest Rate Risk

The direct interest rate risk calculation in Stress Tester 2.0 consists of two parts, reflecting, respectively, flow and stock impacts of interest rate changes. The upper part, in Table D1 works with the repricing gap information from the "Data" worksheet. It calculates the changes in interest income and interest expenses resulting from the "gap" between the flow of interest on the holdings of assets and liabilities in each bucket. The "gap" in each time band or time-to-repricing bucket shows how net interest income will be affected by a given change in interest rates. It sorts assets and liabilities into three time-to-repricing buckets (due in less than 3 months, due in 3 to 6 months, due in 6 to 12 months).[16]

[16] This a very crude breakdown; for more precise calculations, more maturity brackets are utilized, especially for the very short periods, but also for longer periods.

The bottom part of the calculation, in Table D2, shows the impact of interest rate changes on the value of bonds held by the commercial banks. The calculations assume that the bonds are "marked-to-market," i.e. changes in their market value have a direct impact on the capitalization of the banks. The impact of an interest rate change on the market value is approximated using the duration of the bonds held by the banks. The data on duration are given to us by the NBB, which calculated it using more detailed data on the parameters of the bonds and the structure of their holdings by banks—see the formula for duration in the *FSI Compilation Guide* (IMF, 2004).

The direct impact of higher nominal interest rates on capital and capital adequacy is typically negative, resulting from the fact that financial institutions operate with a duration gap between their assets and liabilities. Duration of assets (liabilities), D_A (D_L), is defined as the weighted average, term-to-maturity of an asset's (liability's) cash flow, the weights being the present value of each future cash flow as a percent of the asset's (liability's) full price.[17] Duration approximates the elasticity of the market values of assets and liabilities to the respective rates of return,

$$\frac{\Delta A(r_A)}{A(r_A)} \cong \frac{-D_A \Delta r_A}{(1+r_A)}, \quad \frac{\Delta L(r_L)}{L(r_L)} \cong \frac{-D_L \Delta r_L}{(1+r_L)}, \tag{1}$$

where $A(r_A)$ and $L(r_L)$ are market values of assets and liabilities of the financial system, and r_A and r_L are annual interest rates of assets and liabilities.[18] Differentiating the capital adequacy ratio with respect to the interest rate on assets and substituting from (1), we obtain:

$$\frac{\Delta[C(r_A, r_L)/A_{RW}(r_A)]}{\Delta r_A} \cong -\frac{(L/A_{RW})}{1+r_A}\left(D_A - D_L \frac{1+r_A}{1+r_L}\frac{\Delta r_L}{\Delta r_A}\right)\frac{1-\dfrac{\Delta A_{RW}}{A_{RW}}\dfrac{C}{\Delta C}}{1-\dfrac{\Delta A}{A}\dfrac{C}{\Delta C}}. \tag{2}$$

Assuming that the risk-weighted assets move proportionately to total assets, i.e., $\Delta A_{RW}/A_{RW}=\Delta A/A$, equation (2) can be simplified into

[17] See the *FSI Compilation Guide* (IMF, 2004, paragraph 3.52) for a formula. In practice, the calculation of the duration of total assets and total liabilities of a financial system is a difficult computational task and various simplifications are used in practice (e.g., duration is computed for groups of assets and liabilities with common features and aggregated across such groups; or duration is replaced by residual maturity and time to repricing).

[18] See, e.g., Bierwag (1987). The results are first-order approximations; for large changes in interest rates, second derivative terms need to be included to account for convexity of portfolios. Alternatively, the elasticity of bond prices to interest rate changes can be empirically estimated using past data.

$$\frac{\Delta[C(r_A, r_L)/A_{RW}(r_A)]}{\Delta r_A} \cong -\frac{(L/A_{RW})}{1+r_A} GAP_D , \tag{3}$$

where GAP_D is the duration gap, defined as[19]

$$GAP_D = D_A - D_L \frac{1+r_A}{1+r_L} \frac{\Delta r_L}{\Delta r_A} . \tag{4}$$

Most financial institutions, and banks in particular, operate by transforming short-term, low-interest rate liabilities into long-term, higher-interest rate assets. This means that $D_A >> D_L$, $r_A > r_L$, and $GAP_D > 0$. Thus, an increase in interest rates has a negative impact on the institutions' net worth and capitalization, leading to increased financial sector vulnerability.

In the Bankistan example illustrated in Stress Tester 2.0, only two bonds are available to banks. Both of them are government-issued bonds. The "Assumptions" worksheet shows the key parameters of these bonds (in rows 57–59), and calculates their duration, using the "Duration" function in Excel (see cells H58 and H59).[20]

B. Indirect Interest Rate Risk

The indirect effects, related to the interest-risk nexus work in the same direction. An increase in nominal interest rates—to the extent that it increases real interest rates and makes it more difficult for borrowers to repay their debts and to obtain new credit—is likely to have a negative effect on the credit risk of the financial institutions' borrowers. Other things being equal, higher risk eventually translates into higher losses and a decline in the financial institutions' net worth. The exact impact depends on factors such as the borrowers' earnings in relation to interest and principal expenses, loan loss provisions, and the degree of collateralization of the loans. Country case studies find a positive relationship between higher interest rates and nonperforming loans or loan losses.

The basic calculation does not cover the impact of nominal interest rate changes on real interest rates and thereby on the creditworthiness and ability to repay of the borrowers. The size of the impact depends most importantly on corporate sector leverage and on its exposure to the real estate market (which is also likely to be influenced by changes in interest rates). In order to assess this type of risk, one would usually need to estimate the impact of changes in interest rates on non-performing loans, using a regression model. In our example, the joint

[19] If the interest rates for assets and liabilities move simultaneously, the duration gap can be approximated as a difference of the two durations, $D_A – D_L$.

[20] Excel also has a "Price" function, which can also be used to implement this calculation.

impact of changes in interest rates and in credit quality can be simulated using the worksheet "Scenarios" (see Section IX).

VI. Foreign Exchange Risk

The foreign exchange risk is the risk that exchange rate changes affect the local currency value of financial institutions' assets, liabilities, and off-balance sheet items. The foreign exchange risk is composed of three types: the direct solvency risk (resulting from banks' net open positions in foreign currency and those in local currency that are indexed to exchange rates); the indirect solvency risk (resulting from the impact of foreign exchange positions taken by borrowers on their creditworthiness and ability to repay, and thereby on financial institutions); and the foreign exchange liquidity risk (resulting from liquidity mismatches in foreign currency). In this section, we will focus on the direct solvency risk, and we will discuss implementation of the indirect solvency risk; we relegate the foreign exchange liquidity risk to the section dealing with the liquidity stress test.[21]

The foreign exchange shock in the "FX Risk" worksheet is composed of two parts. The first part, shown in Table E1, tests direct foreign exchange rate risk. The second part, shown in Table E2, shows the impact of the change in the nominal exchange rate on banks through changes in the credit risk.

A. Direct Foreign Exchange Risk

Table E1 in the "FX Risk" worksheet tests direct foreign exchange rate risk based on the net open position in foreign exchange at end-2005. This figure is calculated by the NBB, using the methodology described in the *FSI Compilation Guide* (IMF, 2004), and is copied to Table E1 using a link to the "Data" worksheet.

The *direct exchange rate risk* can be assessed using the net open position in foreign exchange, one of the "core FSIs," defined in IMF (2004). The direct exchange rate risk is arguably the easiest part of stress tests to implement. To illustrate this test, let F denote the net open position in foreign exchange, C the capital, A_{RW} the risk-weighted assets (all in domestic currency units), and e the exchange rate in units of foreign currency per unit of domestic currency. A depreciation (decline) in the exchange rate leads to a proportional decline in the domestic currency value of the net open position, i.e. $\Delta e/e=\Delta F/F$ (for $F\neq 0$). Let us **assume** that this translates directly into a decline in capital, i.e. $\Delta C/\Delta F=1$.[22] The impact of the exchange rate shock on the ratio of capital to risk-weighted assets would then be

---

[21] The direct and indirect solvency risks in foreign exchange are often also referred to as "exchange rate risks."

[22] An alternative, and arguably more realistic, approach is to deduct the impact first from profits (if any), and then from capital. See Section II for a more detailed discussion of this issue.

$$\frac{\Delta[C(e)/A_{RW}(e)]}{\Delta e} \cong \frac{\frac{F}{e}A_{RW} - C\frac{\Delta A_{RW}}{\Delta C}\frac{F}{e}}{A_{RW}{}^{2}} \cong \frac{1}{e}\frac{F}{C}\frac{C}{A_{RW}}\left(1-\frac{\Delta A_{RW}}{\Delta C}\frac{C}{A_{RW}}\right), \qquad (5)$$

which uses the fact that $\Delta C/\Delta e=\Delta F/\Delta e=F/e$. The symbol "≅" means that the equation is only approximate for larger than infinitesimal changes. Equation (1) can be rewritten as

$$\Delta[C(e)/A_{RW}(e)] \cong \frac{\Delta e}{e}\frac{F}{C}\frac{C}{A_{RW}}\left(1-\frac{\Delta A_{RW}}{\Delta C}\frac{C}{A_{RW}}\right). \qquad (6)$$

The term $\Delta A_{RW}/\Delta C$ can have values from 0 to 1, reflecting the degree of co-movement of capital and the risk weighted assets.[23] In the special case of $\Delta A_{RW}/\Delta C=0$, i.e. if the risk-weighted assets do not change, the change in the capital adequacy ratio equals the exchange rate shock times the exposure, measured as a product of the net open position to capital (F/C) and capital adequacy (C/A_{RW}), both of which are "core FSIs" as defined by IMF (2004). This is sometimes used as a short-hand calculation of the direct exchange rate stress test. It should be noted that (6) holds only as a linear approximation, which works well in non-sophisticated financial systems. However, if financial institutions have large positions in foreign exchange options, the relationship between the exchange rate change and the impact on capital can become highly non-linear. In such cases, stress tests based on detailed decomposition of financial institutions' open positions are a superior analytical tool.[24]

In Stress Tester 2.0, a depreciation of the Bankistan dollar against the U.S. dollar from the official exchange rate of 55 B$/US$ to the currently prevailing parallel market rate of 85 B$/US$ is **assumed** (at constant cross exchange rates to the US$). Information on the banks' foreign exchange exposure is provided in Table E1. A depreciation will benefit banks that have a long (positive) open position in foreign currency and hurt banks that have a short (negative) position in foreign currency.

Only a very limited number of banks have short positions, therefore the direct depreciation effects are very small—some banks would even gain from a depreciation. Given that most central banks impose limits on foreign exchange positions to capital, this result is not unusual. For most banking systems, the direct foreign exchange solvency risk is rather small. Banks' net open positions in foreign exchange are typically under close scrutiny from banks' risk managers and supervisors. Banks in some countries have explicit limits on these positions as a percent of the bank's capital (the ceilings typically being in the range of 10–

---

[23] Empirically, $\Delta A_{RW}/\Delta C$ could be estimated by a regression.

[24] As a general point, stress tests should include all relevant off-balance sheet items.

20 percent of capital). In other countries, this is addressed by including the net open positions in the capital adequacy calculation. In general, the open positions tend to be rather small and consequently the direct impact of an exchange rate depreciation (or appreciation) tends to be rather small.

B. Indirect Foreign Exchange Risk

Besides direct depreciation effects, a change in the exchange rate would also influence the creditworthiness and ability to repay of the corporate sector. A change in the exchange rate influences the corporate sector in two main ways: first, it changes its competitiveness relative to the foreign corporate sector; second, it influences the corporate balance sheets directly via firms' net open positions in foreign currencies (for instance, companies can borrow massively in foreign currencies.)

The indirect foreign exchange risk seems to be very important. FSAP missions have generally not been able to collect comprehensive data on the corporate sector's foreign exchange exposure, but those FSAP missions that analyzed the corporate sector in detail generally found that the banking sector's indirect exchange rate risk was more important than its direct one. The indirect foreign exchange rate risk appears to be particularly substantial in countries with closely managed exchange rate pegs

To illustrate the significance of the indirect risk in overall banking sector risk, let us denote the corporate sector's debt, equity, and open foreign exchange position as $D_c(e)$, $E_c(e)$, and $F_c(e)$, respectively.[25] Let us **assume** that, similar to the case of banks' net open position, a percentage change in the exchange rate will translate into the same percentage change in the domestic currency value of the net open position, which will in turn lead to an equivalent change in the corporate sector's equity, i.e., $\Delta E_c/\Delta e = \Delta F_c/\Delta e = F/e$. The impact of the exchange rate on the corporate leverage (D_c/E_c) is then given by

$$\frac{\Delta[D_c(e)/E_c(e)]}{\Delta e} \cong \frac{\frac{\Delta D_c}{\Delta E_c}\frac{F_c}{e}E_c - D_c\frac{F_c}{e}}{E_c^{\ 2}} \cong -\frac{1}{e}\frac{F_c}{E_c}\left(\frac{D_c}{E_c} - \frac{\Delta D_c}{\Delta E_c}\right). \qquad (7)$$

Thus, if the corporate sector is short in foreign exchange, a depreciation (decline) in the exchange rate would lead to an increase in its leverage. Corporate leverage is typically positively correlated with the share of banks' nonperforming loans in total loans (denoted as

[25] Given the practical difficulties involved in obtaining empirical data on open positions in the household sector, we refer here for simplicity only to the corporate sector, even though the theoretical analysis would be essentially the same even if we included the household sector.

NPL/*TL*), i.e., $\Delta(NPL/TL)/\Delta(D_c/E_c)=a>0$.[26] The impact of a change in the exchange rate on the *NPL*/*TL* ratio can then be expressed as

$$\Delta(NPL/TL) \cong a\,\Delta[D_c(e)/E_c(e)] \cong -\frac{\Delta e}{e}\frac{F_c}{E_c}a\left(\frac{D_c}{E_c}-\frac{\Delta D_c}{\Delta E_c}\right). \qquad (8)$$

In the special case when $\Delta D_c/\Delta E_c=0$, the change in the *NPL*/*TL* ratio would equal the exchange rate change times the respective FSI (the net open position), times the parameter *a*, which can be estimated empirically, as shown for example in IMF (2003) or Boss and others (2004). To find the impact on capital adequacy, we can **assume**, as has been done in some FSAP missions, that the credit shock moves some of the previously performing loans into the nonperforming category. By differentiating C/A_{RW} with respect to *NPL*/*TL*, and substituting for NPL/TL from (8), we obtain

$$\Delta(C/A_{RW}) \cong \frac{\Delta e}{e}\frac{TL}{A_{RW}}\left(1-\frac{C}{A_{RW}}\frac{\Delta A_{RW}}{\Delta C}\right)\pi\frac{F_c}{E_c}a\left(\frac{D_c}{E_c}-\frac{\Delta D_c}{\Delta E_c}\right), \qquad (9)$$

where we **assume** (as several FSAP missions have done) that provisions are expressed as a fixed percentage (π) of nonperforming loans, and that they are deducted directly from capital.

The incorporation of the indirect effect makes the analysis of foreign exchange rate risk more complex and dependent on additional assumptions or regression analysis. One of the reasons adding to the complexity of the indirect exchange rate stress test is the fact that it includes the effects on stocks as well as on flows. The calculation of the indirect effect as per (9) would need to reflect the impact of exchange rate changes on the net present value of the corporate sector, which means to take into account changes in the net present value of future earnings. For example, in export-oriented companies, a depreciation could be generally expected to increase their future earnings. In terms of the net present value, the effect would be essentially equivalent to the impact of a long position in foreign currency. However, it may be more practical to calculate the impact on flows, by estimating the elasticity of earnings to interest and principal expenses (an encouraged FSI) with respect to the exchange rate, and then to estimate the relationship between this FSI and the NPL/TL ratio. Alternatively, it would be useful to compile an indicator measuring the corporate sector's flow exposure, e.g., a ratio of foreign exchange earnings to total earnings, or (ideally) a ratio of earnings in foreign exchange to interest and principal expenses in foreign exchange.

[26] IMF (2003) shows that for a panel of 47 countries, a 10 percentage point rise in the corporate leverage was associated with 1.1 percentage point rise in the ratio of NPLs to total loans after a one year lag.

Table E2 in the "FX Risk" worksheet shows the impact of the change in the nominal exchange rate on banks through changes in the credit risk. The impact is approximated here by **assuming** that the change in the NPLs is proportional to the volume of foreign exchange loans in a bank. The idea behind this assumption is that a depreciation would increase the domestic currency value of these loans, which would make it more difficult for the borrowers to repay—for example because some of the loans are extended to borrowers with limited access to foreign currency.

VII. INTERBANK (SOLVENCY) CONTAGION RISK

We have so far **assumed** that there is no contagion among banks in the event of a failure. This assumption is relaxed in the worksheet "Interbank," which presents a basic calculation of the interbank contagion risk.[27]

In this section, we focus on contagion through insolvencies ("interbank solvency contagion risk"). There is also, potentially very important, the risk of liquidity contagion through bank runs triggered by a run on another bank. A basic model of liquidity contagion is shown as part of the liquidity risk (Section VIII). The principle of having a matrix of bank-to-bank "exposures" is the same in all the contagion tests, but the specification of the matrix is different for a liquidity test.

We focus here on contagion within the banks in the (domestic) banking system. The Bankistan banking system system includes foreign-owned banks, but we do not study cross-border exposures and cross-border contagion. In principle, the same framework would have to be applied, only the definition of "system" would have to be wider to include the foreign banks. For simplicity (and because obtaining good bank-by-bank data on cross-border exposures is not trivial in practice), we focus here on interbank exposures in the domestic market. We have to be at least aware, however, that in addition to the domestic exposures, there may also be important cross-border exposures and the related risk of contagion.

The upper panel in the "Interbank" worksheet (Table F1) derives a matrix of interbank exposures for the twelve banks in Bankistan. The first panel shows the matrix of net interbank credits, in which each cell shows, for the bank in the column (i.e., for the bank listed in the same column in row 4) its net credit to the bank in the row (i.e. the bank listed in the same row in column A). For example, the value of 70 in cell I11 means that the bank in column I (i.e., DB1) has an outstanding net credit to the bank in row 11 (i.e., SB2) in the amount of B$70 million. A corresponding entry of -70 in cell G13 confirms this by showing

[27] In a real stress testing exercise, we would have to start from the very beginning with balance sheets that include the interbank exposures. This would make the aggregation of the bank-by-bank data more cumbersome (we would have to net out the bank-to-bank exposures). For simplicity, this exercise starts without the interbank exposures, which are added only later.

that the bank in column G (i.e., SB2) has an outstanding negative net credit of B$70 million to the bank in row 13 (i.e., DB1), i.e. it has a net borrowing of the same amount. The matrix is obtained by netting out the gross figures in the interbank lending table reported in the "Data" worksheet, i.e. reported by the NBB. In this matrix, positive figures mean that the bank in the column is a net creditor of the bank in the row, while negative figures mean that the bank in the column is a net borrower of the bank in the row. The diagonal cells are left empty, since the focus of this exercise is on exposures between each bank and the other banks.

To facilitate further calculation, the matrix of net interbank credit is converted into the matrix of net interbank exposures. This is done by "stripping" the matrix of net creditors by focusing on the positive numbers (because those are the banks exposed to interbank credit risk). The other cells in the table are left empty. For example, the value of 70 in cell I26 means that the bank in column I (i.e., DB1) has an outstanding net exposure to the bank in row 26 (i.e., SB2) in the amount of B$70 million. The corresponding cell, G28, is empty: the bank in column G (i.e., SB2) has no net exposure to the bank in row 28 (i.e., DB1), since SB2 is a net borrower with respect to DB1, and therefore has no direct credit exposure to DB1.

A. "Pure" Interbank Contagion

The middle panel in the "Interbank" worksheet (Table F2) provides a calculation of the "pure" interbank contagion exercise. The exercise shows what would happen with the capital of the bank in the column if the bank in the row failed and defaulted on all its interbank borrowing. This is actually a series of 12 separate stress tests, one in each row, showing for each bank what would be the direct impact of its failure on the capital of each of the other banks. The stress test is run in several iterations, as the contagion-induced failures ("first iteration") can induce failures in other banks ("second iteration"), which can lead to further failures ("third iteration"), and so on.

The first part of Table F2 shows, in each row, the post-shock capital of the individual banks in the column after the **assumed** failure of the bank in the row. For example, row 49 shows in columns F to Q what would be the post-shock capital of banks SB1 to FB4 if bank SB3 failed. The results in row 49 suggest that none of the banks would have a negative post-shock capital as a direct result of SB3's failure.

The next part of Table F2 shows which banks fail as a result of the first iteration ("1" denotes a newly failed bank, all others have zeros). The table shows that two banks would fail as a result of failures in other banks, namely DB1 can fail as a result of failures in SB2 or FB3 (see cells I62 and I71) and DB2 can fail as a result of failures in SB1, SB2, FB1, and FB3 (see cells J61, J62, J69, and J71).

We are **assuming** for simplicity that if a bank's capital stays positive after an iteration, the banks does not fail and remains able to repay all its interbank obligations; if its capital becomes negative, it fails and does not repay its obligations. The calculation can be made more realistic by estimating a more complex mapping between the capital adequacy ratio and the bank's probability of failure (rows 6 and 22 in the "Assumptions" worksheet contain an example of such a mapping). Such mapping is likely to indicate a wider scope for interbank contagion than the introductory calculation presented here; however, the calculations would be based on similar techniques.

To keep the calculations straightforward, we **assume** here that the impact of shocks is deducted directly from capital. It is not complicated, however, to extend the file to take into account banks profits. We include this as one of the possible tasks in Appendix I.

The second iteration needs to be calculated only for the failures in SB1, SB2, FB1, and FB3, because these are the only banks whose failures lead to failures in other banks (namely, DB1 or DB2). The failures in DB1 and DB2 could lead, in the second iteration, to failures in other banks if the other banks had substantial net credit outstanding with respect to DB1 or DB2. However, as illustrated in rows 73–113, this is not the case: the banks that are creditors to SB3 and DB1 have exposures to SB3 and DB1 that are well below their capital.

Rows 73–85 starts the second iteration by inverting the table of the failed banks from the first iteration. This is to illustrate that in the second iteration, the banks that were exposed to contagion in the first iteration become the source of further contagion.[28] Rows 88–99 show the capital after the second iteration, which generally equals the capital after the first iteration, but for failures in SB1, SB2, FB1, and FB3 (i.e. in rows 88, 89, 96, and 98), the second iteration capital of the bank in the column is lowered by the amount of exposure of this bank to the bank that failed in the first round. This is highlighted in rows 101–112, which show the change in capital between the first and second iteration: capital is lower for stress tests in rows 101, 102, 109, and 111, because the corresponding banks (SB1, SB2, FB1, and FB3) cause additional bank failures in the second iteration.

Rows 115–126 show that the third iteration is not needed in this case, as none of the banks are affected enough to fail in the second round. If such failure occurred, the third round (and any subsequent round) could be implemented in the same way as the second round.

[28] The calculation can in principle be implemented without the inversion. However, inverting the table makes calculations easier, because we can use the SUMPRODUCT function. Also, the inversion highlights the nature of the calculation: in the second iteration, we look at the failed creditors from the previous iteration, and analyze who are their creditors. Since the exposure table has creditors in columns and borrowers in rows, going from the first iteration to the second one requires inverting either the matrix of exposures, or the table with results from the first iteration. In this implementation, we chose the latter approach.

To present the results in terms of capital adequacy, the worksheet uses an **assumption** on the risk-weight of the interbank loans that become unpaid. Reflecting the typical Basel weight for interbank loans, this assumption is set at 20 percent, but can be changed by changing the value of the B77 cell in the "Assumptions" worksheet.

The "pure" interbank contagion test could be interpreted as a measure of systemic importance of individual banks: the bigger the decline in the system's capital (or capital adequacy ratio), the more systemically important is the bank whose default is assumed. In our example, FB2 is the systemically most important bank, using this criterion, because assuming its failure yields the lowest post-shock systemic capital (B$ 3,570 million in the cell B138) and post-contagion CAR (9.7 percent in the cell B152). This test (or rather set of tests) is useful, but it does not take into account the different likelihood of failures in different banks, an issue that is addressed by the "macro" interbank contagion test.[29]

B. "Macro" Interbank Contagion

The lower table in the "Interbank" worksheet (Table F3) shows a "macro" interbank contagion test. In this case, we model the case when bank failures are triggered by macroeconomic developments, in particular by the scenario that is already modeled in the "Scenario" sheet (Figure 4). The starting point for the "macro" interbank contagion is therefore the post-shock values of capital and risk-weighted assets for each bank from the "Scenario" sheet.[30] For the failed banks (SB2 and SB3 for the default settings in the spreadsheet), we run an interbank contagion exercise using the matrix of net interbank exposures. Then, we search for banks that fail in this first iteration. If there were no new failures, the contagion exercise would stop here. In our example, however, there are three new failures: SB1, DB1, and DB2. We therefore run a second iteration, looking at the impact of these additional failures on other banks. We find that they lead to one additional failure, namely in DB4. If this failure led to other new failures, we would need to run a fourth iteration. However, in this particular case, the contagion-induced failure of DB4 does not lead to other failures, and the process stops at the third iteration.

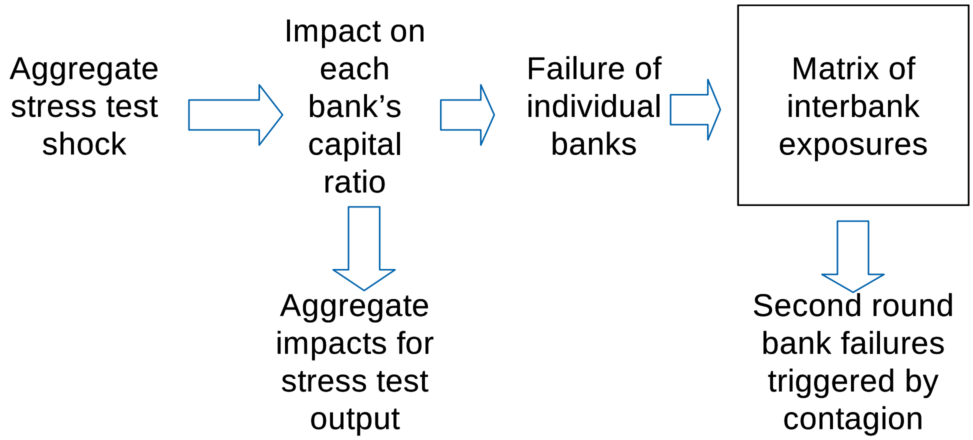
What is the key difference between the "pure" and the "macro" contagion tests? The "pure" contagion test **assumes** that a failure occurs in a single bank, for example for some internal reason (e.g., because of a large fraud in the bank); it does not distinguish the relative

[29] The "pure" interbank contagion tests are more common in the literature. For example, Sveriges Riksbank presents their results regularly in its Financial Stability Report. "Macro" interbank contagion tests are less common, but are presented for example by Elsinger, Lehar, and Summer (2003) for Austria or by Čihák, Heřmánek, and Hlaváček (2007) for the Czech Republic.

[30] The calculations in the "Scenario" sheet disregarded these contagion effects, so the calculation in Table F3 of the "Interbank" sheet can be seen as an extension of the calculation in the "Scenario" sheet.

likelihood of the failure of various banks. This is what the "macro" contagion test does. It analyzes situations when all banks are weakened at the same time by a common external (typically macroeconomic) shock, which affects each bank differently depending on its exposures to the various risk factors, and makes some of the banks (perhaps more than one) fail. For the default settings for Bankistan, for example, the first iteration of the "macro" contagion test involves a simultaneous failure of two banks (SB2 and SB3), resulting in a second iteration failure of three others banks (SB1, DB1, and DB2), leading to a third iteration failure of DB4. The process stops at the third iteration, which does not lead to additional failures.

Figure 4. 'Macro' Interbank Contagion



VIII. Liquidity Tests and Liquidity Contagion

Testing for liquidity risks is less common in central banks' stability reports and in IMF work than testing for risks to solvency (see Appendices II and III, respectively). This reflects mostly the fact that modeling liquidity risks is more complex. First, to properly model liquidity fluctuations in banks, one needs to have very detailed, high frequency data, such that are typically used by commercial banks themselves in their liquidity management models. Second, to model the impact of large liquidity shocks, one needs to consider the broader liquidity management framework, in particular the lender of last resort function of many central banks.

At the same time, testing for liquidity risks is important. In the last two decades, much of the attention in risk management and prudential supervision was on capital, partly in relation to the efforts to standardize capital adequacy requirements across countries. In the process, relatively less attention has been paid to cash flows and analysis of liquidity (e.g.,

Goodhart, 2006). Analyzing the response of liquidity to stress is an important undertaking, because liquidity is how a stressful situation often manifests itself in the short run.

The presentation of the stress test impact is different from the solvency tests discussed so far. The impact is shown for each bank in terms of the number of days it would be able to survive a liquidity drain *without resorting to liquidity from outside* (other banks or the central bank).[31] This is a relatively narrow approach to liquidity stress testing, but it is one that allows for an introductory exposition without going into details.

The Stress Tester 2.0 file contains two basic examples of liquidity tests. The "Liquidity" worksheet contains these two examples in two tables. Figure 5 shows the results of these two tests for the default values contained in the spreadsheet.

Table G1 models a liquidity drain that affects all banks in the system proportionally, depending on their volumes of demand and time deposits. The worksheet allows the user to change **assumptions** on the percentage of demand deposits and time deposits that get withdrawn each day and on the percentage of liquid assets and other assets that banks can convert to cash each day.[32]

Table G2 models "liquidity contagion," where the liquidity drain starts in the smallest or weakest banks and we test how this can affect the larger or stronger banks. The test allows for **three possible measures** of "bank safety": (i) total assets; (ii) total assets, with a premium for state ownership; and (iii) pre-shock rating. In the first case, depositors perceive bank safety as linked to the size of the bank, approximated by total assets. In the second case, they also perceive state-owned banks safer than privately-owned banks (because of an explicit or implicit government guarantee in the former case). In the third case, depositors' perceptions of bank safety are correlated with the banks' recent financial performance.[33] Table G2 also combines the liquidity impact of government default with a bank run. It allows users to change the **assumption** on the percentage of the government bonds which are in default.

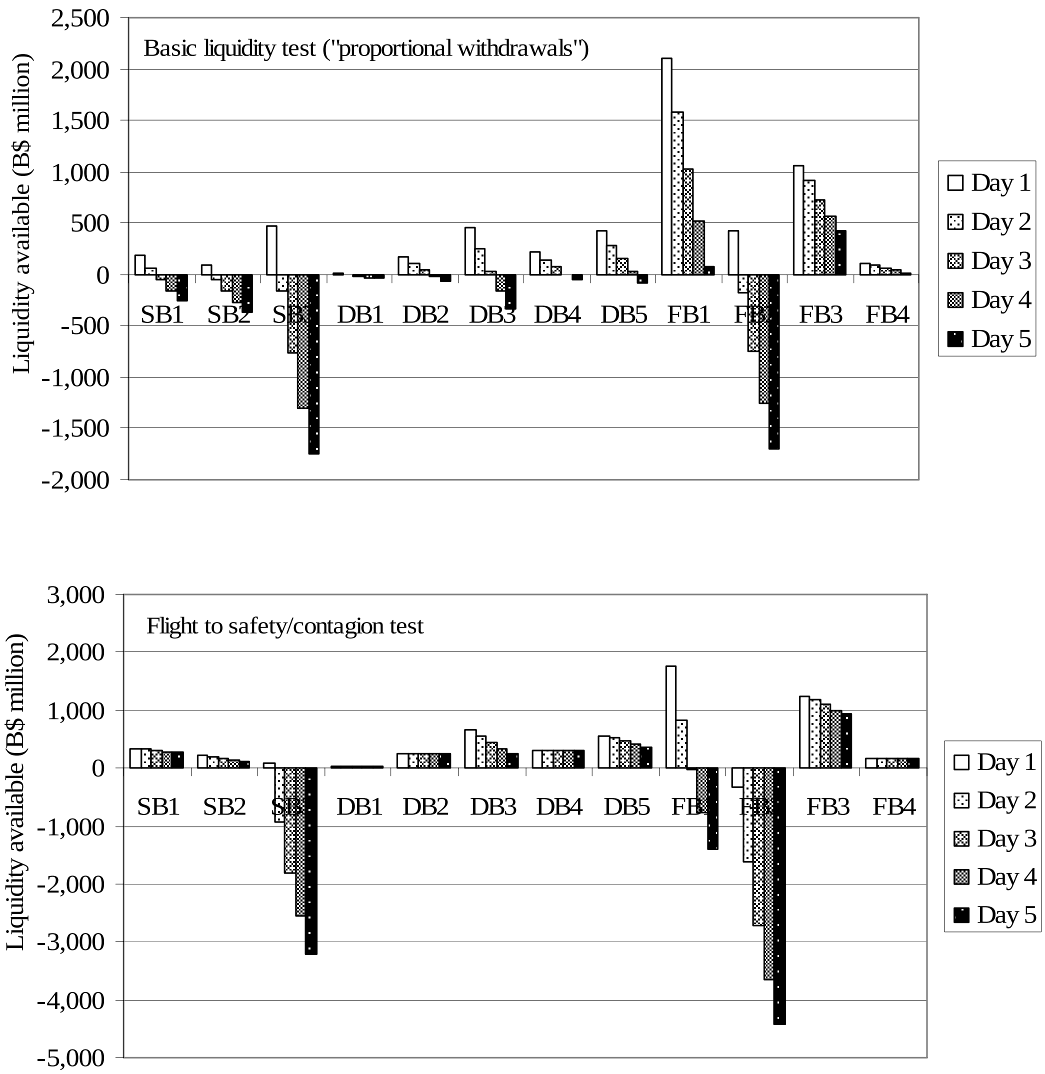
---

[31] As a rule of thumb, some supervisors see 5 days as an important threshold for a bank's ability to withstand a liquidity run. The number 5 is typically chosen because after 5 days or less, banks will close for a weekend or a holiday, giving some "breathing time" for the bank management and supervisors to regroup, assess the situation, and decide on measures and public announcements to make. Of course, this traditional rule of thumb has been partly diluted with the growth of direct (and in particular internet) banking.

[32] Another useful distinction in many cases is between domestic and foreign exchange deposits (and assets). It is possible to extend the file to assume different withdrawal rates by currency of denomination (we relegate this extension as an issue for Appendix I).

[33] Liquidity contagion can also be thought of in terms of a similar "exposure" matrix as used in the solvency contagion, only instead of "net uncollateralized interbank exposures," the value of the cell in the matrix for liquidity contagion would be the difference between bank *i*'s and bank *j*'s measure of "bank safety."

Figure 5. Results of Liquidity Stress Tests



IX. Scenarios

The “Scenarios” worksheet illustrates, in Tables H1–4, how shocks to the various risk factors can be combined into a single scenario. The main reason for using scenarios rather than single factor shocks is that in the macroeconomic context, changes in several risk factors are typically interrelated. For example, a large increase in nominal interest rates can lead to an increase in real interest rates, which can (perhaps with a lag) contribute to an increase in NPLs. If this is the case, banks will be hit not only by the direct impact of the nominal increase in interest rates, but also by the indirect impact through credit risk.

The purpose of this worksheet is for the users to see how different combinations of the shocks impact the capital adequacy of the system. For simplicity of presentation, the

worksheet contains only formulas linked to the other worksheets, namely "Credit Risk" for credit risk, "Interest Risk" for the interest rate risk, "FX risk" for the direct and indirect foreign exchange solvency risk, "Interbank" for interbank contagion, and "Liquidity" for liquidity risk. When a number of different specifications of a stress test are available, Stress Tester 2.0 allows the user to specify which of the alternative specifications is being considered in the overall scenario. These choices are **assumptions** that can be made in cells B69–75 of the "Assumptions" worksheet (a summary of the chosen scenario is then shown on top of Table H in the "Scenarios" worksheet). The "Scenarios" worksheet adds up the impacts of the selected shocks to arrive at an aggregate impact (see Box 5 for a discussion of the appropriateness of adding up the individual impacts).

As a basic presentational approach, the "Scenarios" worksheet shows the impacts of the scenarios in terms of the capital adequacy, and decomposes the overall impact into the individual risk factors (in percentage points of the CAR ratio). The charts in the "Assumptions" worksheet allow the user to immediately and graphically see the results of the various assumptions on the outcome in terms of changes in capital adequacy ratios. For illustration, we reproduce these charts here as Figure 6. The values shown in these charts reflect the default sizes of shocks and starting assumptions in the accompanying file. Black indicates the baseline (pre-shock) values of the capital adequacy ratio, red indicates the corresponding post-shock values, and yellow indicates values after the shocks and after taking into account the subsequent contagion among banks. The pre-shock (baseline) CAR values are taken from the "Data" worksheet, while the values corresponding to the stressful scenario are taken from the "Scenario" worksheet.

The worksheet also compares the overall impact with the banks' profits. While we have so far **assumed** that the impacts are deducted directly from capital, in reality banks could use profits as their first line of defense. Table H in the "Scenarios" worksheet shows for comparison, for each bank, what its profits were in the past. It also allows the user to **assume** (in the appropriate green cell) an autonomous shock to net interest income.

A. Designing Consistent Scenarios[34]

How can one design consistent scenarios? In general, there are two ways of asking questions about exposures in the financial system. The first way is to ask, for a given level of plausibility, which scenario has the worst impact on the system ("the worst case approach"). The second way is to ask, for a given impact on the system, what is the most plausible combination of shocks that would need to occur to have that impact ("threshold approach").

[34] For more details, see Čihák (2004a, 2005).

Box 5. Can We Add The Impacts of Shocks?

The “Scenarios” worksheet combines the shocks from the other worksheets. Can the impacts be added up?

First, it is important to take into account concentration of risks in institutions. Simply adding up aggregate losses caused by individual shocks could overlook situations when risks are concentrated in an institution or a group of institutions. Stress Tester 2.0 addresses this issue by calculating the impacts bank-by-bank. This allows us to see whether some banks are hit by the selected combination of shocks much harder than others (they indeed are).

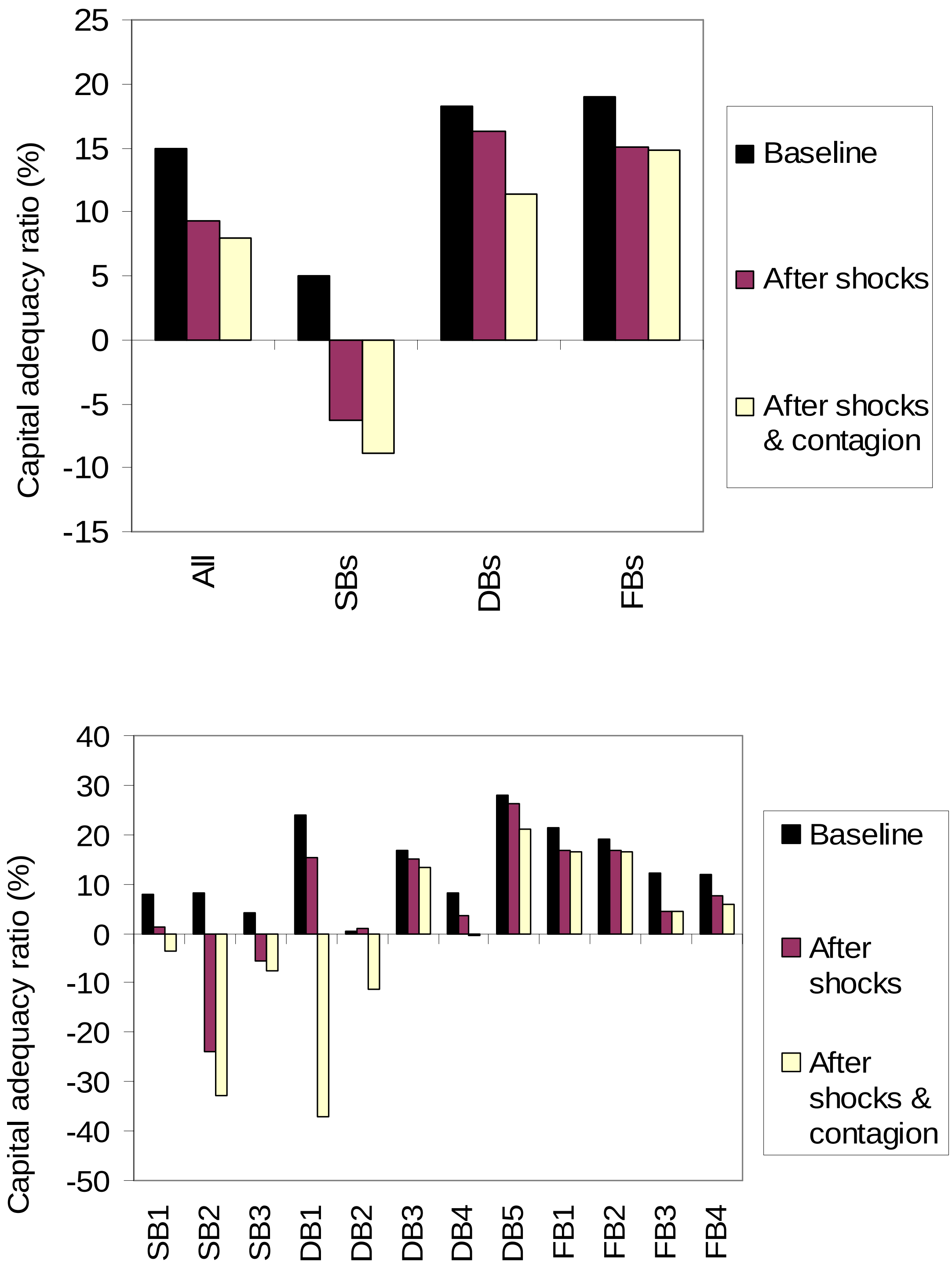
Second, it is not trivial to combine solvency and liquidity risks. Stress Tester 2.0 illustrates how this can be done, at least to some extent. It uses the NBB’s supervisory early warning system to combine the changes in solvency and liquidity (and other measures) to identify the change in the supervisory rating, and the implied change in probability of default. There are some limitations to this approach (in particular, the PDs cannot be easily aggregated for the system as a whole), but it provides a useful illustration.

Figure 7 shows the process of scenario selection under the worst case approach and the threshold approach, for a simplified case when there are only two risk factors (for example, changes in interest rates and exchange rates). Each ellipse depicts the set of combinations of the two risk factors with the same probability of occurrence. The shape of the ellipse reflects the correlation between the two factors, and its size reflects the level of plausibility (the larger the ellipse, the smaller the plausibility). The diagonal lines depict combinations of the risk factors leading to the same overall impact, measured here by a change in the system’s capital adequacy ratio (CAR). The impact increases with the size of the shocks to the risk factors, so the CAR decreases in the northeast direction. The diagonal lines do not have to be straight; they are only depicted here as such for simplicity. Figure 7 illustrates that the worst case approach and the threshold approach are two essentially equivalent ways of analyzing the same problem.[35]

The worst case approach starts with selecting a level of plausibility (e.g., 1 percent), and searching for the combination of shocks with this level of plausibility that have the worst impact on the portfolio. This means searching for the point on the largest ellipse that lies as far northeast as possible. In Figure 7, this is point A.

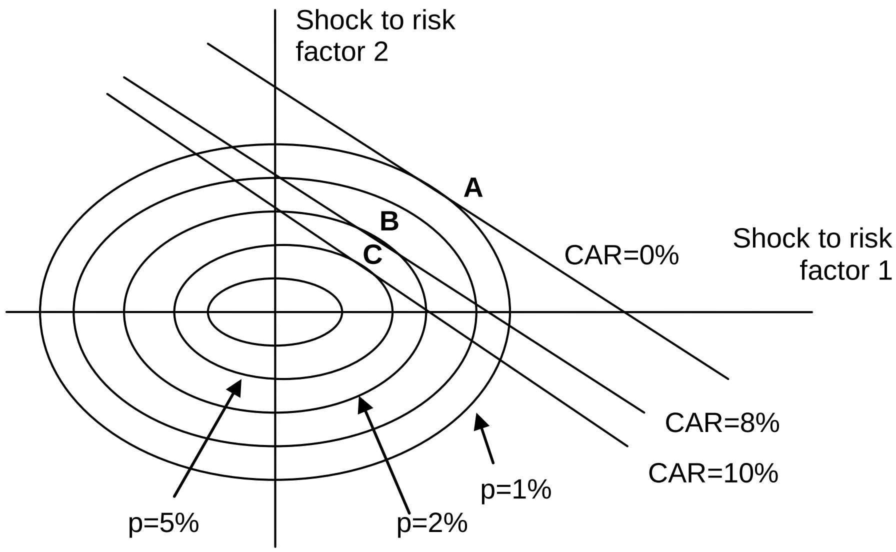
[35] To some readers, these two approaches may resemble the dual tasks of microeconomics.

Figure 6. Impact of Stress on Capital Adequacy Ratios



Source: Stress Tester 2.0.xls, "Assumptions" worksheet, based on default values of shocks and assumptions.

Figure 7. 'Worst Case Approach' vs. 'Threshold Approach'



The threshold approach starts with selecting the threshold, i.e. the diagonal line; it then searches for the most plausible (i.e., smallest) shocks reaching this threshold. This is straightforward if there is only one risk factor; if there are two risk factors, one needs to take into account the correlation between the risk factors. For the specific correlation pattern in Figure 7, selecting a threshold of zero capital adequacy would lead again to the combination of shocks corresponding to point A.

Establishing the plausibility level of a scenario can be difficult in practice, given that the scenario should be a low probability, "tail" event. For risk factors with good time series of historical data (in particular, for market risks), the natural starting point is to base the scenarios on the past volatility and covariance patterns. Calibrating the shocks is particularly straightforward for single-factor stress tests: an exchange rate shock can be based on 3 standard deviations of past exchange rate changes (corresponding roughly to a 1 percent confidence level). With multiple risk factors, one needs also to look at the covariance statistics of the variables, or use stochastic simulations based on macroeconomic models. Such calculations are subject to a number of caveats. In particular, models can break down for large shocks. Nonetheless, the models, if used cautiously, can help to find a first-cut approximation of stress test scenarios (see Box 6 for an additional discussion of choosing the "right" scenario).

Box 6. Picking the 'Right' Scenario

In discussions on designing stress tests, too much is often made of establishing the "right" scenario. Of course, it is important, at least in theory, for scenarios to be internally consistent, as highlighted, for example, by Jones, Hilbers, and Slack (2004) or by Figure 7 in this paper. In practice, assessing such consistency is tricky, because the scenarios are also supposed to be exceptional (but plausible).

How to address this challenge? One approach is to choose a concrete extreme historical scenario (e.g., the East Asian crisis of 1997) and calculate what would be the impact of repeating such a scenario (or an adaptation of such scenario) in the present situation of the banking system. The main advantage is that historical scenarios are easy to communicate and to implement. Also, they are plausible, because such a situation actually happened. Their main disadvantage is that past crises may not be good models for future crises. Also, the probability level of the past historical scenario may be unclear.

Another approach, used by some FSAPs and central bank FSRs, is to use an existing macroeconomic model (e.g., a model used by the central bank for macroeconomic forecasts and policy analysis) as a basis for stochastic simulations showing the distribution of the key risk factors in the case of shocks to the model's exogenous variables. The challenge of this approach often arises from the fact that such macroeconomic models typically do not include a measure of credit risk (e.g., nonperforming loans to total loans, or another measure of asset quality). Therefore, this approach usually involves estimating a "satellite" model that links a measure of credit risk to the variables from the macroeconomic model. Unlike the macroeconomic model, the satellite model can be estimated (and generally should be, if adequate data area available) on individual bank (and even individual borrower) data. The estimates from the satellite model can then be used for the balance sheet implementation.

Yet another, and perhaps more direct, approach can be to plot the existing observations of the various risk factors (in a similar fashion as shown, for two risk factors, in Figure 7) against a measure of soundness (e.g., capital adequacy ratio) and use this to identify the most stressful combinations of risk factors (in terms of Figure 7, this would mean identifying the points lying the most towards the north-east).

In sum, there is a range of methods, each with its advantages and disadvantages. Picking a scenario that is stressful and tells an interesting and consistent story is important. However, in most cases, identifying "the right scenario" is an equivalent of finding the Holy Grail. Much more important than fine-tuning scenarios is (i) being transparent about the underlying assumptions of the scenarios; (ii) being transparent about the sensitivity of the results to those assumptions; and (iii) showing how results with the same assumptions change over time. Showing results over time allows making judgments about the developments in the overall pool of risks and in the structure of risks faced by a financial system.

B. Linking Stress Tests to Rankings and Probabilities of Default

The "Scenarios" sheet also illustrates the links between the stress tests and the supervisory early warning system. It follows the same approach as the illustration shown in the "Data"

worksheet, only it is derived from the bank-by-bank data after the shocks rather than those before the shocks. Specifically, Table H2 provides the post-shock FSIs and other ratios for individual banks. These can be compared with the corresponding pre-shock ratios, shown in Table A3 (in the "Data" worksheet).

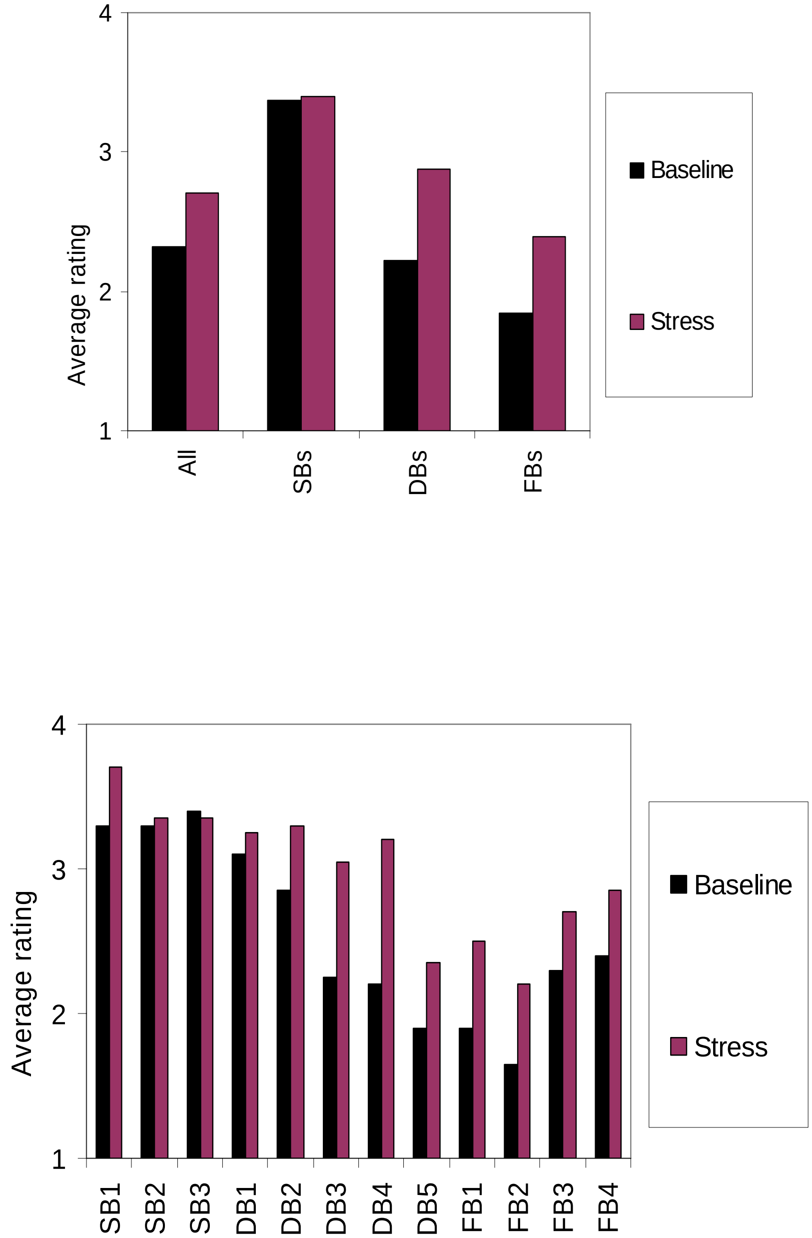
Table H3 converts these ratios into post-shock ratings of the individual banks, using the same "step functions" that were applied in Table A5 to the pre-shock ratios. In both cases, the "step function" reflects the off-site supervisory assessment model, and is specified in the "Assumptions" worksheet (rows 3-22). Table H3 also provides averages (weighted by banks' total assets) for the three peer groups and for the banking system as a whole.

Table H4 converts the post-shock ratings into post-shock probabilities of default for each bank. These can be compared with the pre-shock probabilities of default, shown in Table A6 in the "Data" worksheet.

The charts in the "Assumptions" worksheet, reproduced in this paper as Figure 8 and Figure 9, illustrate such comparisons. Black indicates the baseline values of the indicators (ratings in Figure 8 and probabilities of default in Figure 9), and red indicates the corresponding values in situations of stress. The values of the baseline are taken from the "Data" worksheet, while the values corresponding to the stressful scenario are taken from the "Scenario" worksheet. The values shown in the charts reproduced here reflect the default sizes of shocks and starting assumptions in the accompanying file.

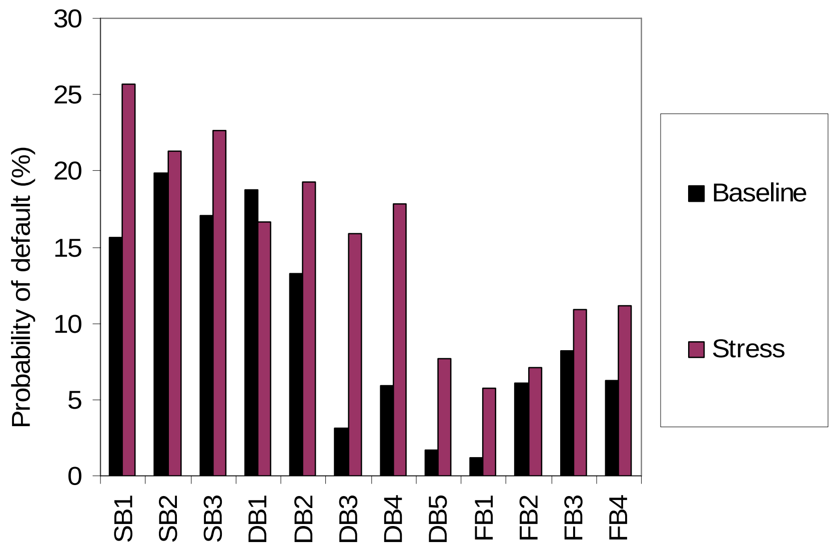
Figure 10 shows another form of presentation of stress testing results. It captures the impact of stress on the banks' z-scores. As indicated earlier, the z-score has become a popular measure of bank soundness, because it is directly related to the probability of a bank's insolvency. Figure 10 illustrates how the banks' z-scores decline as a result of the assumed impacts, generally mirroring the decline in banks' probabilities of default, shown in Figure 9 (one needs to bear in mind that higher z-scores correspond to lower probabilities of default). If the user changes the key sizes of shocks and assumptions in the accompanying Excel file, the charts in the file will change automatically.

Figure 8. Impact of Stress on Supervisory Ratings



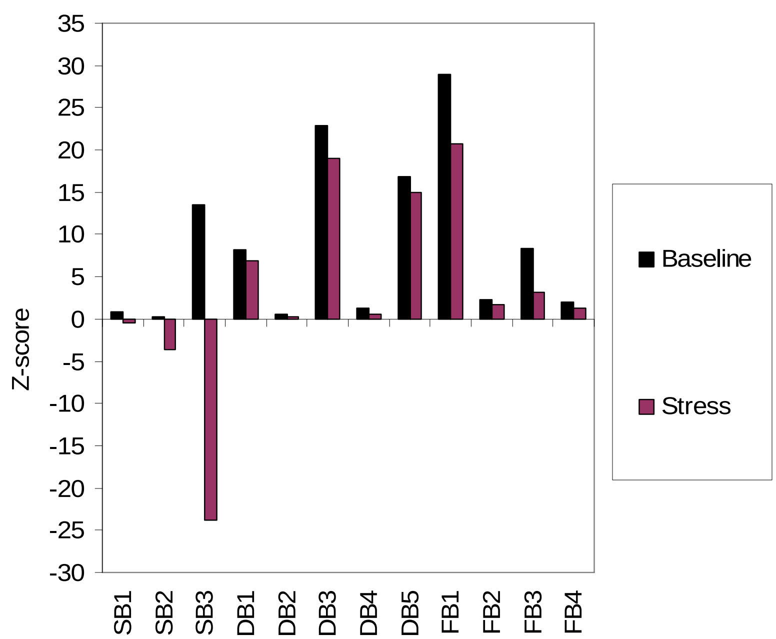
Source: Stress Tester 2.0.xls, “Assumptions” worksheet, based on default values of s hocks and assumptions.

Figure 9. Impact of Stress on Banks' Probabilities of Default



Source: Stress Tester 2.0.xls, "Assumptions" worksheet, based on default values of shocks and assumptions.

Figure 10. Impact of Stress on Banks' Z-Scores



Source: Stress Tester 2.0.xls, "Assumptions" worksheet, based on default values of shocks and assumptions.

C. Modeling the Feedback Effects

The stress test calculations presented here focus on the impacts of shocks arising from the macroeconomic environment and impacting the financial sector. From a macroeconomic perspective, an important issue is whether the shocks in the financial sector can have feedback effects impacting the macroeconomic environment.

The practical problem in modeling the feedback effects is that there are too many. One direct effect that can be easily incorporated into the financial programming framework employed by the Fund, is an impact on capital on "Other Items Net" in the monetary survey and thereby on other macroeconomic variables.

However, this is just one of many potential impacts. In some cases, the effects depend on the behavior of the institutions in situations of stress. For example, if banks attempt to sell off certain types of assets (e.g., real estate) in situations of stress, they may bring down the asset prices, with repercussion effects for other sectors (e.g., household consumption). Also, bank failures triggered by stress may result in a credit crunch.

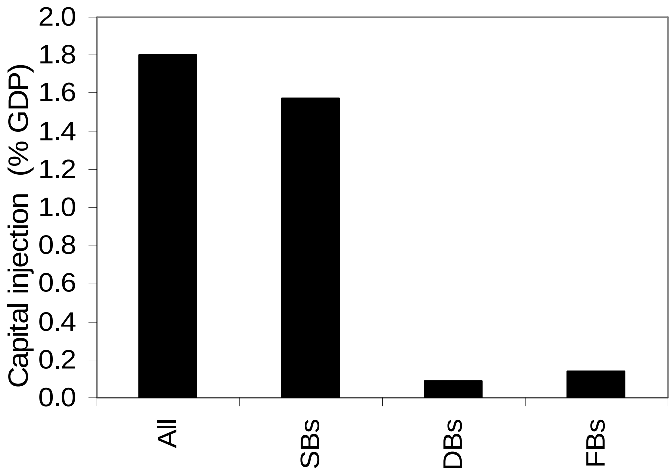
In the accompanying Excel example, we approximate the potential macroeconomic impacts by the capital injection needed to bring all banks to the minimum required capital adequacy ratio. This indicator does not capture all the potential macroeconomic effects, but it is a useful broad indicator of potential fiscal costs associated with averting failures in the banking system. It is an upper-bound estimate of such costs. The public sector will most likely inject capital into state-owned banks, but it is less clear whether (and if so, how much) it will inject into the other banks, which are privately owned (it is generally more likely to do so for larger banks that are considered "too big to fail"). For this reason, the charts and results in the file show a breakdown of the necessary capital injection by ownership. The charts are included in the "Assumptions" worksheet, and are reproduced here as Figure 11. The values shown in the figure reflect the default sizes of shocks and default assumptions in the accompanying file.

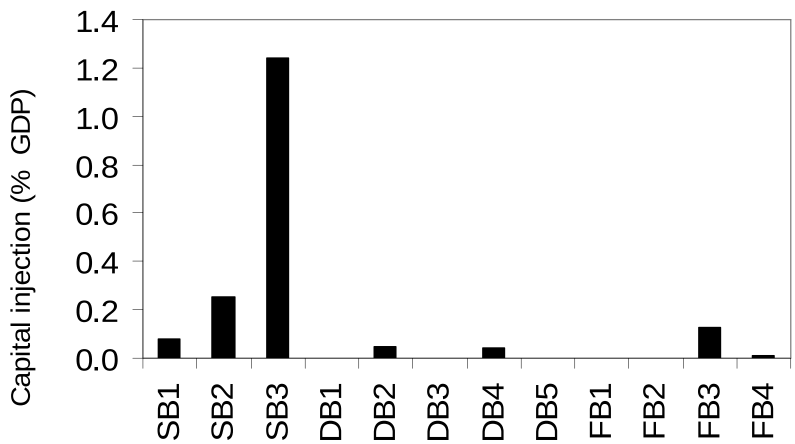
X. CONCLUSIONS AND EXTENSIONS

It is difficult for a person to really understand stress testing until he or she goes through actual stress testing calculations. This exercise enables readers to do just that: they can change the assumptions and observe changes in results, or they can even try to develop the file further.

Stress tests are complementary to other tools for financial stability analysis, and the exercise illustrates this complementarity. In particular, it illustrates that stress tests are complementary to the FSIs, which allow for "benchmarking," i.e. for a baseline assessment of the financial

Figure 11. Capital Injections Needed to Bring Banks to Minimum Capital Adequacy





Source: Stress Tester 2.0.xls, "Assumptions" worksheet, based on default values of shocks and assumptions.

system under no stress. FSIs can also be used to describe the impact of stress on a system. Stress tests are also complementary to the supervisory early warning system, an example of which was shown as well. The early warning system is traditionally used to calculate ratings and probabilities of default in the "baseline" scenario, but as shown in Stress Tester 2.0, it can also be used to produce ratings and probabilities of default in a stressful scenario.

In addition, stress tests are complementary to other tools that have not been illustrated in the accompanying file, even though they were briefly mentioned. Those include assessments of compliance with regulatory standards and codes (e.g., Basel Core Principles or Insurance

Core Principles). Also important is an assessment of the broader financial stability framework, which includes financial safety nets (liquidity support mechanisms, deposit protection and other protection schemes, and crisis management).

In the exercise, we have tried to highlight several main challenges, which should be borne in mind by anyone attempting stress testing:

- Stress testing is *data intensive*. Stress tests deal with low-probability events. This implies that there will always be a lack of data for stress tests. Carrying out stress tests therefore almost by implication often requires simplifying assumptions. The key thing an analyst should do is to be transparent about the nature of the assumptions.
- To make things worse, *nonlinearities* are likely to kick in for the large shocks that are being contemplated in stress tests.
- Macroeconomic and other *models tend to break down* in crises. Past crises may not be a good guide for the future. As an example, a change in borrower characteristics may leave credit more vulnerable to interest rate risk.
- The impact of shocks is *distributed over time*. It takes time for asset quality to deteriorate and for that deterioration to have an impact. When a system finds itself in a crisis, that crisis will evolve over a period of time, sometimes many years. Modeling stressful scenarios therefore has to take the time dimension into account (for example, the Bank of England's stress tests look at impacts over a period of 2–3 years). It is important to be clear what the benchmark scenario is, against which we compare the stressful scenario.
- A number of *mitigating measures* can be taken by participants and authorities, especially if we are looking at longer time periods. Also, feedback effects start playing a role over time.

Two main steps can be taken to address these challenges:

- *Keep assumptions transparent, and be clear on the sensitivity of the results to the assumptions.* It is important to be transparent about both the underlying assumptions of the scenarios as well as the sensitivity of the results to those assumptions. The accompanying stress testing file is trying to do just that. All assumptions are brought into one place and highlighted. Users can experiment with the assumptions to see the impact of the changes on the results.
- *Present stress test results over time.* Presenting results over time helps to say whether the overall pool of risks has changed, or whether the structure of risks has changed. Most financial stability reports still do not provide results over time, which makes interpretation of the presented results more difficult for readers (Čihák, 2006). The Stress

Tester 2.0 exercise is based on a one-period snapshot, but the idea behind the exercise is that it is run repeatedly, replacing the input data (in yellow) by data from other periods.

The idea behind the accompanying file is that it can be developed in a modular fashion, with additional modules capturing additional risks or elaborating on the existing ones. One can think about Stress Tester 2.0 as a module in a broader stress testing toolkit, and about the cells with assumptions as interfaces between this module and other modules. The following is a list of extensions that are worth considering:

- *Credit risk-macro nexus.* Credit risk is the main source of risk in most financial systems. At the same time, it is the part where this exercise is perhaps the most simplified. Ideally, we would need to have more detailed data on loan exposures and loan performance by economic sector, and also data on the financial soundness of the corporate and household sectors. Time series of historical data should ideally be used to establish linkages between macroeconomic variables and loan performance (or, in the absence of reliable time series, estimates from other countries could be used). The calculations presented in this exercise try to convey the gist of the credit risk stress test, without going into technical details. Box 4 provided references to studies that analyzed this nexus in more detail.

- *Credit VaR models.* Related to the previous point, developed country commercial banks and some banks in medium-income countries are using value-at-risk (VaR) models as a basis for stress testing for credit risk. There is a range of these models, all of which share a common purpose of determining the probability distribution of the losses on a portfolio of loans and other debt instruments. Being able to compute the loss distribution of a portfolio is critical, because it allows the determination of the economic capital required by credit operations. Implementing VaR tests for credit risk on the macroprudential level is more complex than for market risk VaR models, because (i) less banks use them; and (ii) the risk factors and their parametrization is likely to differ across banks, making comparability and aggregation more difficult. Nonetheless, implementing credit risk models on a macroprudential level is possible and can provide a useful benchmark for credit risk evaluation, as illustrated in Avesani and others (2006) on the CreditRisk+ model. Indeed, some recent FSAPs have made use of these approaches (Appendix III).

- *Stress-testing based on factor models*. This includes: (i) portfolio risk management models using structural credit risk models of obligors assets and risky debt based on domestic and international factor models, such as the KMV portfolio manager models, CreditMetrics, or the default and conditional probability of default models (e.g., Segoviano and Padilla, 2006); (ii) Merton-type structural models of banks which calibrate risk adjusted balance sheets and implied assets of banks which are then linked to domestic and international factor models. Stress testing can then analyze how changes in key international and domestic factors drive individual bank risk and systemic risk).

These types of models can be applied to banking systems in 30 or so middle and some low income countries and to banking systems in developed countries.

- *Stress testing models that use the a combined structural bank model with an interest rate term structure model*. These models, such as the Shimko-Tejima-van Deventer model for bank soundness and capital adequacy, try to assess the impact of interest rate level, interest rate volatility, and asset-interest rate correlations on individual bank risk and systemic risk. (e.g.,Belmont, 2004).
- *Other risk factors.* Depending on the sophistication of the financial system and on the type of its exposures, it may be necessary to perform also stress tests for other risk factors. These might include asset price shocks (including,e.g.,shocks to real estate prices), and shocks to commodity prices (especially in developing countries with significant exposures to commodities).

APPENDIX I. QUESTIONS FOR THE HANDS-ON EXERCISE

This appendix lists questions that can be used in a hands-on seminar based on the attached Excel file. The questions can be used to gauge whether seminar participants (or readers) have become familiar with the concepts introduced in the paper.

Credit risk

Credit shock 1 ("Underprovisioning")

(1) Which banks would be undercapitalized or even insolvent after applying the provisions?

(2) How much capital would the government need to inject in order to bring the capital asset ratio of the undercapitalized/insolvent banks up to the required minimum of 10 percent?

(3) Discuss the assumptions underlying this calculation and ways in which the calculation can be improved. What if we find that the provisioning rates are inappropriate? For example, assume that we find that the general provisioning rates in neighboring or similar countries are 2 percent for pass loans and 5 percent for special mention loans rather the numbers used in Bankistan (1 percent and 3 percent, respectively). How do the results differ if we adjust the domestic provisioning rates for the difference? (Hint: change the appropriate cells in the "Credit Risk" sheet).

Credit shock 2 ("Increase in NPLs")

(4) Which banks can not withstand a 25 percent increase in NPLs and will be undercapitalized or fail in this event?

(5) How much capital would the government need to inject in order to bring the capital asset ratio of the undercapitalized/insolvent banks up to the required minimum of 10 percent?

(5) What happens if only assets with a risk-weight of 0 percent, i.e. government bonds, would be affected?

(6) Discuss alternative assumptions regarding the increase in NPLs and how these can be implemented.

Credit shock 3 ("Sectoral Shocks")

(7) Which banks are the most exposed to shocks to tourism? Which ones are most exposed to shocks to agriculture?

(8) How can this methodology be improved? What other breakdowns of loans could be used?

Credit shock 4 ("Large Exposures")

(9) Which banks are the most exposed to concentration risk?

(10) How can this exercise be used to model risks from non-bank financial institutions?

Interest rate risk

(10) Which banks can withstand the increase in interest rates and which would fail?

(11) What are the associated potential costs for the government given the failure of banks in times of stress?

(12) What are the weaknesses of the interest rate risk calculations and how can they be remedied?

Foreign exchange risk

Direct FX risk

On the basis of a depreciation rate of 55 percent, answer the following questions:

(13) Which banks can withstand the shock and which would fail or be undercapitalized?

(14) What are the associated potential costs for the government given the failure of banks in times of stress?

Indirect FX risk

Looking at Table E2, answer the following questions:

(15) How does the inclusion of the indirect FX risk change the result of this stress tests?

(16) How can the incorporation of the indirect FX risk be made even more realistic?

Liquidity risk

(17) How would you model a case when there are different withdrawal rates for domestic currency and foreign currency deposits?

(18) What other variables can be used to approximate depositor's perceptions of banks' safety? How would you model it?

(19) How would you make the test for government default more realistic?

Interbank contagion

(20) How would you calculate the impact of the third iteration in the "pure" contagion exercise?

(21) How would you incorporate banks' profits into the exercise?

(22) How would you reflect in the exercise the fact that banks with positive, but small, capital adequacy ratios can also fail?

(23) How would you calculate the impact of the third iteration in the "macro" contagion exercise, if there were a need for the third iteration?

Creating consistent scenarios

(24) Discuss the consistency of the shocks that are currently assumed in the worksheet for the three risk factors. Is this a reasonable, i.e. extreme but plausible, scenario, given what you know about the macroeconomic vulnerabilities in Bankistan? What adjustments to the scenario are needed?

(25) How do the various peer groups differ in their vulnerability to the various risk factors? How are the results sensitive with respect to the various assumptions? How are the results sensitive with respect to various assumptions on profits and shocks to the net interest income?

APPENDIX II. STRESS TESTING IN FINANCIAL STABILITY REPORTS

Stress testing has become a widely used analytical tool in the analysis of financial stability. The share of central banks' financial stability reports (FSRs) that included summaries of stress tests was about 75 percent as of the end of 2005 (Čihák, 2006). A common pattern is that in early FSRs, stress tests are not used; they are included only in subsequent reports after the central bank becomes more comfortable with the basic presentation. This indicates that the importance of stress testing as an analytical tool in financial stability work is growing as more central banks seek to improve the quality of their FSRs.[36]

Appendix Table 1 gives an overview of stress tests in FSRs in general, and Appendix Table 2 provides concrete country examples. Stress tests presented in most FSRs are relatively rudimentary, as central banks are still only at the beginning of their work in this area. As an important caveat, it should be recognized that a vulnerability can be analyzed in some cases even without conducting a formal stress test. For example, even when an FSR does not contain an explicit stress test for exchange rate risk, the report would typically include a discussion of the open positions in foreign currency. Similarly, when a central bank does not publish an FSR or does not include stress tests in its FSR, it may still carry out stress tests for internal purposes, without publishing the results.

There is substantial cross-country variation in the size and range of shocks covered, and in the methodologies applied. This reflects the differences in the financial systems, the risks that they face, as well as differences in the quality and structure of the available data. Nonetheless, the stress tests in FSRs show several common features:

- *Stress tests in FSRs tend to have a wide coverage of the banking sector*, including either all banks or virtually all in terms of market share. Other parts of the financial sector are covered much less, even though there are exceptions (for example, the German FSR includes stress tests for the insurance sector).

- *A majority of presented stress tests are based on bank-by-bank data.* This can be understood as a recognition that stress tests done on aggregate data can miss some potentially important risks arising from the concentration of risks in weaker institutions. Central banks that are not involved in microprudential supervision (e.g., Bank of England or Bank of Norway) and do not have access to supervisory data are

[36] The ratio of FSRs publishing stress test results declined in 2004–2005. This may be only a reflection of the rapid growth in new FSRs: central banks often start publishing FSRs without stress tests and only after they become more comfortable with the publication, they start including stress test results (Čihák, 2006).

more likely to rely on top-down approaches or calculations based on non-supervisory data.

- *Credit risk is covered in almost all stress tests.* Interest rate risk is covered in most stress tests. Exchange rate risk is covered in some, but in many cases it is analyzed only in terms of open positions, without an explicit stress test.

- *Most stress tests presented in FSRs are sensitivity calculations.* Some include scenario analysis, based on historical or hypothetical scenarios. Only a few stress tests are based on an econometric model. When models are used, they tend to be relatively rudimentary compared to those used in other central bank work, such as inflation forecasting. Inclusion of indirect exchange rate effects and contagion is rare. When the latter is done, it is a basic exercise based on net interbank market exposures.

- *Virtually all the surveyed stress tests in FSRs have been positive in the overall assessment of the financial sector, suggesting that it is stable.* This has mirrored the fact that FSRs tend to have overall positive findings (see Čihák, 2006). Stress tests in FSRs tend to confirm the overall positive conclusions by finding that the system is robust, and capable of withstanding substantial shocks.

- *In many cases, the interest in stress tests was spurred by a recent FSAP mission.* In some cases, recent stability report(s) included a summary of the FSAP stress tests (e.g., Austria and Netherlands), and it is likely that the central bank will continue with its own stress testing program, broadly along the FSAP lines. In other cases (e.g., Denmark and Norway), the FSR included "FSAP-style" stress tests even before the country had undergone an FSAP.

Appendix Table 1. Stress Testing in FSRs: Overview, End of 2005

Topic	Percent of FSRs
Stress testing included	55
Stress testing follows a recent FSAP	38
Credit risk stress testing included	55
Interest rate risk stress testing included	45
Exchange rate risk stress testing included	33
Other risks included	33
Scenario analysis included	38
Contagion analysis included	10
Credit risk based on an econometric model	8

Source: author's calculations based on individual FSRs.

Even though a majority of FSRs now include stress tests, there is still a number that do not, and most of those that do could give them more prominence and present them on a comparable basis over time. As regards the coverage of the stress tests, most of them would benefit from a broader scope that would cover the credit risk, its interplay with market risk, and interbank contagion. Stress tests could be linked to a broader macroeconomic scenario. If such a set of tests were run regularly with the same assumptions, it would allow the reader to see changes over time in the overall pool of risks and in the structure of risks faced by the financial system. Presenting results as a single point in time is interesting, but does not allow the reader to appreciate the developments in the overall pool of risks or changes in the structure of risks. It would therefore be particularly useful to compare current stress test results with those shown in the past. The specific improvements in stress testing that can be recommended in many FSRs include the following:

- *Making greater use of scenarios.* Most FSRs that include stress tests define them very narrowly; thus the tests typically comprise only a set of single factor shocks. In those FSRs that include scenarios testing, the scenarios are not well justified; it is not clear how the design of the scenarios relates to recent history or future risks.

- *Integrating the impacts and the buffers.* Results of the stress tests are often presented in terms of specific loan loss provisions. This does not take into account how well the risks and exposures are matched by buffers (profits and capital). This may result in biased results, depending on whether risks and exposures are concentrated in weakly capitalized institutions or in well capitalized institutions. Presenting impacts in terms of capital (or capital adequacy or profitability) allows a better assessment of the concentration of risks. To carry out this type of analysis, it is important to have data for individual institutions.

- *Placing a greater focus on liquidity tests.* The stress tests presented in FSRs typically focus on testing for the impact on banks' solvency. It is almost as important to test for liquidity; however, explicit liquidity tests are quite rare in FSRs.

- *Adding a contagion analysis.* In most FSRs that present stress test results, the analysis can be further improved by analyzing contagion among banks and also between nonbanks and banks. To carry out this type of analysis, it is important to have data for individual institutions. The interbank contagion calculation can include two broad approaches. The first one focuses on the risk of insolvency through the interbank market. An important requirement for these calculations is a matrix of net uncollateralized interbank exposures. The calculations can be implemented in two basic ways: (i) assume a failure in one institution (e.g., because of mismanagement); or (ii) run a macro-related contagion test, where the first round of failures ("fundamental failures") is triggered by a macroeconomic stress test scenario, and

then a contagion is run through the system to see whether this leads to another round (and perhaps even more rounds) of failures. The second broad approach focuses on the risk of liquidity runs. An important requirement for this approach is detailed data on withdrawals in past episodes of bank runs.

- *Making use of stressful scenarios.* Bowen, O'Brien, and Steigum (2003) noted, in an otherwise very positive review of the Norges Bank's FSR, that the report tended to rely on a relatively narrow and mild set of assumptions (e.g., by assuming that banks would be able to earn a satisfactory level of earnings even in periods of stress). Also, a more general review of FSRs suggests that for those countries that have had an FSAP and publish FSRs, the scenarios presented in the FSRs tend to be generally less stressful than those presented in FSAPs. Assessing the plausibility of a scenario is a complex task; as a minimum, one can recommend that the assumptions are kept consistent across the FSRs (to allow for comparability and to avoid using more optimistic assumptions in periods of weakness) and that FSRs include also an approach to defining the scenarios that does not start from the plausibility of a scenario.

- *Employing a threshold approach.* The prevalent approach to stress testing starts by presenting scenarios with a probability that is unknown (at least to the reader) and not easy to calculate. In some FSRs, it may be useful to employ also a "threshold approach," which instead starts by asking what shock would it take to make the system reach a certain threshold (e.g., reaching the system's capital adequacy of 8 percent or making a certain percentage of institutions insolvent). This approach was used in several FSAPs and FSRs (e.g., the National Bank of Poland's Financial Stability Review).

Appendix Table 2. Examples of Stress Tests in Financial Stability Reports 1/

Country	*Coverage*	*Main conclusion*	*Credit shock*	*Interest rate shock*	*Exchange rate shock*	*Other shock*	*Scenario*	*Indirect FX risk*	*Contagion*
Austria 2/	All banks	Stress testing supported positive assessment of banks' risk-bearing capacity.	Incr. in loan loss provisions to loans by 30 percent.	Upward shifts in EUR, USD, CHF curves; downward shift in YEN curve.	Appreciation/depreciation of EUR by 10 percent. Worst case estimation.	Equity price risk	Yes	Yes	Yes
Czech Republic	All banks	System basically stable. Recent shift from credit risk to interest rate risk.	Increase in NPLs by 30 percent or in the NPL/TL ratio by 3 percentage points.	Increase by 1 pct point/2pct points. Combination of weighted gap and duration methods.	Domestic currency depreciates by 15 or by 20 percent.		Yes, two scenarios	Yes	Yes
Denmark	6 Nordic groups and 46 (out of 99) Danish institutions.	The banking institutions increased their resilience considerably.	An increase in losses on loans to non-public sector by 1 or 2.25 percentage points.	An increase in interest rates by 1 or by 3 pct points. Decrease in the average lending rate by 1 pct point.		Decrease in stock prices by 30 %. Decrease in net fee and commission income by 40 %.	Yes, several combinations of the shocks listed here.	No	No
Germany	A sample of banks; insurance companies	No risk to financial stability at present.	Credit risk estimated by an econometric model.	Twists of the yield curve at the short end, parallel shifts across all maturities, and fluctuations in the medium-term range.	EUR appreciates or depreciates by 15%.	30 % decline in stock prices in all markets.	Yes, "oil price scenario," int. rate scenario." Credit risk scenario using an econometric model.	No	No
Hungary 3/	All active banks	Results indicate an improvement in the sectors' resilience.	4 shocks to NPLs (e.g., doubling, increase by 2 st. dev.)	Domestic rates: +500 bp, -300 bp Foreign: +/-200 bp	+/- 40 %	No	No	No	No
Latvia	All banks	Vulnerability to overall credit risk decreased in 2004; household lending the biggest risk.	Increase in NPL ratio by 3 percentage points. A number of sectoral shocks, assuming that a percentage of loans to some sectors become NPLs.	No explicit stress test included, the report notes that most loans are floating rate.	USD depreciating by 10 % against EUR.		No.	No	No

Appendix Table 2. Examples of Stress Tests in Financial Stability Reports 1/

Country	*Coverage*	*Main conclusion*	*Credit shock*	*Interest rate shock*	*Exchange rate shock*	*Other shock*	*Scenario*	*Indirect FX risk*	*Contagion*
Poland	All banks	The system exhibits high stability.	Three shocks: (i) satisfactory and special mention loans migrate to doubtful; (ii) substandard and doubtful migrate to loss; and (iii) bankruptcy of three largest borrowers.	Not a stress test, but an analysis of gains/losses on interest-sensitive instruments, and the maturity of debt securities.	Not a stress test, but an analysis of VaR and open positions.	Equity price risk and property market risks analyzed (but without a stress test)		No	No
Nether-lands 2/	Major fin. institutions (84% banks, 54 % insur.c., 50 % pens.f.)	Banks are sufficiently shock-resistant.	+/-50 bps change in credit spreads (larger for insurance and pensions)	+/-100 bps parallel move; 50 bps flattening/steepening of yield curves (larger for ins&pen)	+/-10 % change in the exchange rate of EUR vs. other currencies	+/-15 % change in all relevant stock indices; 25 % increase in market volatilities	Yes, "domestic crisis of confidence, "dollar crisis"	No	Yes
Norway	All banks / seven largest conglome-rates.	Short-term stability outlook satisfactory. However, increased vulnerability of household sector.	Decline in economic growth, increased unemployment.	Interest rates unchanged, but interest burden of real sector increased appreciably.		A fall in property prices reduces mortgage values, causing a rise in loss given default.	Yes, all tied to credit risk.	No	No
Sweden	Four major banks	The major banks improved their potential for coping with shocks.	Failure of the largest counterparty, assumed recovery ratio of 25 percent.	Increase in interest rates by 1 pct points, and a 30 percent fall in the stock market.			No	No	Yes

1/ Abbreviations: NPL ... nonperforming loans, TL ... total loans, pct... percentage, bp ... basis points, st. dev. ... standard deviation, int. rate ... interest rate, EUR ... Euro, USD ... U.S. dollar, CHF ... Swiss Frank, and YEN ... Japanese yen.
2/ The latest FSR contained the stress tests carried out by (or in collaboration with) an FSAP mission.
3/ Based on the FSR at the end 2003. The subsequent two FSRs presented only the "stress CAR," which shows a bank's financial position in a situation where all NPLs are written off.

Source: The author, based on central banks' recent financial stability reports. Austrian National Bank: Financial Stability Report 7, June 2004. Danmarks Nationalbank: Financial Stability 2003. Deutsche Bundesbank: Report on the Stability of the German Financial System, Monthly Report, October 2004. De Nederlandsche Bank: Overview of Financial Stability in the Netherlands, December 2004, Issue No. 1. Hungarian National Bank: Report on Financial Stability, June 2003. National Bank of Poland: Financial Stability Review, First Half of 2004. Norges Bank: Financial Stability, 2004:1, June 2004. Sveriges Riksbank: Financial Stability Report, 2004:2

APPENDIX III. STRESS TESTING IN THE FINANCIAL SECTOR ASSESSMENT PROGRAM

Stress tests are a crucial tool of the quantitative part of the stability assessments in the Financial Sector Assessment Program (FSAP). All FSAP missions have included some form of stress tests, ranging from a very basic assessment of sensitivity of income sources to ad-hoc assumptions, to elaborate exercises involving Monte Carlo simulations or models developed by financial institutions themselves. The tests have been very much tailored to country-specific circumstances, in particular to reflect the complexity of the financial system, and data availability.

The Independent Evaluation Office of the IMF reviewed stress testing in the FSAP as part of its evaluation of the FSAP program (Independent Evaluation Office, 2006). It found that the overall quality of stress testing work was very high, but it noted that the reporting of the stress testing results often takes a “black-box” approach, with too little discussion of limitations implied by data and methodological constraints and choices on which shocks to analyze. It recommended greater “health warnings” about the interpretation of results. It also noted that there is a considerable gap between “good practice” approaches to modeling shocks and those used in many other cases. It also noted that some assessments avoided analyzing the consequences of politically sensitive shocks (e.g., public debt defaults).

Over time, stress tests in FSAPs have become more sophisticated. Older assessments mostly contained a single-factor sensitivity analysis, based on historical extremes. Stress tests in more recent assessments are generally characterized by a greater focus on scenario analysis, greater involvement of the authorities (including macroeconomic models or micro-level credit models), greater involvement of financial institutions (including for internal models and value-at-risk calculations), and more frequent inclusion of interbank contagion and nonbank financial institutions (Appendix Table 3).

Appendix Tables 4–11 provide an overview, for European FSAPs, of some of the basic features of stress tests. The reason for focusing on European FSAPs, apart from keeping the length of the appendix within reasonable bounds, is the fact that Europe is the first region to be completely covered by FSAP missions. For specific examples of European stress tests, see IMF (2005b, for France), IMF (2006, for Spain), Boss and others (2004, for Austria), Hoggarth (2005, for the United Kingdom), and Čihák (2004b, for the Czech Republic).

Appendix Table 3. Evolving Role of Stress Testing in the FSAP, 2000–2005

(Percent of all FSAPs Initiated in the Period)

	2000–2002	2003–05
Scenario analysis	64	95
Contagion analysis	11	38
Insurance sector stress testing	25	37

Source: the author's update of IMF and World Bank (2005a).

Appendix Table 4. Who Did the Calculations in European FSAP Stress Tests? 1/

	FSAP
Supervisory agency/central bank	Austria, Belgium, Denmark, Estonia, France, Germany, Hungary, Ireland, Israel, Malta, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom
FSAP team	Belarus, Belgium, Bosnia and Herzegovina, Croatia, Czech Republic, Denmark, Estonia, Iceland, Ireland, Israel, Latvia, Lithuania, Macedonia, Moldova, Norway, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Serbia, Spain, Ukraine, United Kingdom
Financial institutions	Belgium, Denmark, Estonia, France, Germany, Ireland, Israel, Luxembourg, Malta, Netherlands, Norway, Portugal, Spain, United Kingdom

1/ In some FSAPs, calculations were done by several parties, as indicated in the table.

Appendix Table 5. Institutions Covered in European FSAP Stress Tests

Institutions Covered	**FSAP**
All banks (bank by bank)	Belarus, Belgium, Lithuania, Moldova, Ukraine
Large/systemically important banks (bank by bank)	Austria, Belgium, Bosnia and Herzegovina, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Israel, Latvia, Luxembourg, Malta, Netherlands, Norway, Poland, Romania, Russia, Slovakia, Slovenia, Serbia, Spain, Sweden, Switzerland, United Kingdom
Insurance companies	Belgium, Denmark, Finland, France, Netherlands, Norway, Portugal, Spain, Sweden, United Kingdom
Pension funds	Netherlands, United Kingdom

Appendix Table 6. Approach to Credit Risk Modeling in European FSAP Stress Tests

Approach to Credit Risk Modeling	**FSAP**
NPLs, provisions: historical or macro-regressions	Austria, Czech Republic, France, Iceland, Israel, Russia, Romania, Sweden
NPLs, provisions: ad hoc approaches	Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, France, Hungary, Israel, Latvia, Lithuania, Macedonia, Malta, Moldova, Poland, Serbia, Slovakia, Slovenia, Switzerland, Ukraine
Shocks to probabilities of default based on historical observations or regressions	Austria, Belgium, Denmark, Luxembourg, Russia, Spain
Shocks to probabilities of default (ad hoc)	Germany, Netherlands, Norway, United Kingdom
Explicit analysis of cross-border lending	Austria, Spain
Explicit analysis of foreign exchange lending	Austria, Croatia
Explicit analysis of loan concentration	Malta, Netherlands, Russia, Serbia

Appendix Table 7. Interest Rate Shocks in European FSAP Stress Tests

Interest Rate Scenarios Used	**Examples of Shock Sizes**
• Ad hoc or hypothetical interest rate increase • Parallel shift in yield curve • Historical interest rate increase • Basel Committee Amendment to Capital Accord to incorporate market risk	• 3 standard deviations of 3-month changes • 50%-100% increase • three-fold increase in nominal rate • 100 basis point shock to interest rates • 100 basis point shock to dollar interest rates and a concomitant 300 basis point shock to local currency interest rates • 300 basis point increase

Appendix Table 8. Approaches to Interest Rate Modeling in European FSAP Stress Tests

Approach to Interest Rate Risk Modeling	**FSAP**
Repricing or maturity gap analysis	Austria, Belarus, Belgium, Croatia, Czech Republic, Hungary, Lithuania, Macedonia, Malta, Moldova, Poland, Russia, Romania, Serbia, Ukraine
Duration	Belgium, Iceland, Israel, Latvia, Norway, Slovakia, Switzerland
Value at Risk	France, Denmark, Germany, Israel, Netherlands, United Kingdom
Others (e.g., Δ NPV of balance sheet, Δ market value of bank capital, regressions, simulations)	Norway, Sweden

Appendix Table 9. Exchange Rate Shocks in European FSAP Stress Tests

Exchange Rate Scenarios Used	**Examples of Shock Sizes**
•Ad hoc or hypothetical devaluation • Historical large exchange rate changes	• 20%-50% devaluation • 30% devaluation • 10% depreciation • 20% depreciation/appreciation • 40% depreciation/appreciation of Euro/Dollar exchange rate

Appendix Table 10. Approaches to Exchange Rate Modeling in European FSAP Stress Tests

Approach to Exchange Rate Risk Modeling	**FSAP**
Sensitivity analysis on the net open position	Austria, Belarus, Belgium, Bulgaria, Croatia, Czech Republic, Hungary, Iceland, Latvia, Lithuania, Macedonia, Malta, Moldavia, Norway, Poland, Russia, Romania, Serbia, Slovakia, Slovenia, Sweden, Switzerland, Ukraine
Value at Risk	France, Germany, Israel, Netherlands, United Kingdom

Appendix Table 11. Approaches to Modeling Other Risks in European FSAP Stress Tests

Other Risk Modeling Approaches	**FSAP**
Liquidity risk (ad-hoc decline in liquidity)	Austria, Belarus, Belgium, Bosnia and Herzegovina, Germany, Lithuania, Netherlands, Russia, Spain, Ukraine, United Kingdom
Liquidity risk (historical shock)	Croatia, France
Shock to main stock market index	Austria, Belgium, Finland, France, Germany, Israel, Latvia, Lithuania, Malta, Netherlands, Norway, Slovakia, United Kingdom
Housing price shock	Netherlands, Norway, United Kingdom, Ukraine
LTV ratios, mortgage PDs	Croatia, Sweden
Commodity price	Finland
Interbank contagion	Austria, Belgium, Luxembourg, Netherlands, Romania, United Kingdom
Slowdown in credit growth	Bosnia and Herzegovina
Country risk	Luxembourg
Competition risk (i.r. margin)	Lithuania, Slovenia
Shock to specific sector(s)	Belarus, Finland

References

Avesani, Renzo G. Kexue Liu, Alin Mirestean, and Jean Salvati, 2006, "Review and Implementation of Credit Risk Models of the Financial Sector Assessment Program," IMF Working Paper No. 06/134 (Washington: International Monetary Fund).

Baboucek, Ivan and Martin Jančar, 2005, "Effects of Macroeconomic Shocks to the Quality of the Aggregate Loan Portfolio," Czech National Bank Working Paper No. 1/2005 (Prague: Czech National Bank), available at http://www.cnb.cz/www.cnb.cz/en/research/cnb_wp/download/cnbwp_2005_01.pdf.

Berkowitz, J., 1999, "A Coherent Framework for Stress-Testing," Finance and Economics Discussion Series (Washington: Board of Governors of the Federal Reserve Board).

Belmont, David P., 2004, *Value Added Risk Management in Financial Institutions: Leveraging Basel II & Risk Adjusted Performance Measurement* (Hoboken, NJ: Wiley)

Blaschke, Winfrid, Matthew T. Jones, Giovanni Majnoni, and Soledad Martinez Peria, 2001, "Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences," IMF Working Paper No. 01/88 (Washington: International Monetary Fund).

Blåvarg, M and Niemander, P., 2002, "Inter-bank Exposures and Systemic Risk." Sveriges Bank Economic Review No. 2/2002 (Sweden: Sveriges Riskbank).

Boss, Michael, 2002, "A Macroeconomic Credit Risk Model for Stress Testing the Austrian Credit Portfolio." *Financial Stability Report* No. 4 (Vienna: Austrian National Bank).

Boss, Michael, Gerald Krenn, Markus Schwaiger, and Wolfgang Wegschaider, 2004, "Stress Testing the Austrian Banking System," Austrian National Bank, *Österreichisches Bankarchiv* 11/04, pp. 841–852, available at http://www.oenb.at/de/img/artikel_9_841852_boss_et_al_11_tcm14-27296.pdf.

Bowen, Alex, Mark O'Brien, and Erling Steigum, 2003, "Norges Bank's Financial Stability Report: A Review" (Oslo: Norges Bank), available at http://www.norges-bank.no/front/rapport/en/fs/2003-01/vedlegg.html.

Boyd, John H., and David E. Runkle, 1993, "Size and Performance of Banking Firms," *Journal of Monetary Economics*, Vol. 31, pp. 47–67.

Bunn, P, Cunningham, A and Drehmann, M, 2005, "Stress Testing as A Tool for Assessing Systemic Risks," *Bank of England Financial Stability Review*, June, pp. 116–26.

Crockett, Andrew, 1997, "Why Is Financial Stability a Goal of Public Policy?" In: *Maintaining Financial Stability in a Global Economy, a Symposium* (Jackson Hole, Wyoming: Federal Reserve Bank of Kansas City), available at http://www.kc.frb.org/publicat/econrev/pdf/4q97croc.pdf.

Čihák, Martin, 2004a, “Stress Testing: A Review of Key Concepts,” Czech National Bank Research Policy Note No. 2/2004, available at http://www.cnb.cz/en/pdf/IRPN_2_2004.pdf.

———, 2004b, “Designing Stress Tests for the Czech Banking System.” Czech National Bank Research Policy Note No. 3/2004, available at http://www.cnb.cz/en/pdf/IRPN_3_2004.pdf.

———, 2005, “Stress Testing of Banking Systems,” *Finance a úvěr/Czech Journal of Economics and Finance*, Vol. 55, No. 9–10.

———, 2006, “How Do Central Banks Write on Financial Stability?“ IMF Working Paper No. 06/163 (Washington: International Monetary Fund), available at www.imf.org/external/pubs/ft/wp/2006/wp06163.pdf.

———, 2007, ”Systemic Loss: A Measure of Financial Stability,” *Finance a úvěr/Czech Journal of Economics and Finance*, forthcoming.

———, and Jaroslav Heřmánek, 2005, “Stress Testing the Czech Banking System: Where Are We? Where Are We Going?” CNB Research Policy Note, 2/2005, available at www.cnb.cz/en/research/research_publications/irpn/download/rpn_2_2005.pdf

———, Jaroslav Heřmánek, and Michal Hlaváček, 2007, “New Approaches to Stress Testing the Czech Banking Sector,” *Finance a úvěr/Czech Journal of Economics and Finance*, forthcoming.

Committee on the Global Financial System (CGFS), 2005, Stress Testing at Major Financial Institutions: Survey Results and Practice,” Report by a working group established by the CGFS (Basel: Bank for International Settlements), available at www.bis.org.

Danmarks Nationalbank, 2004, “Market-Based Risk Measures for Banks,” in *Financial Stability 2004* (Copenhagen: Denmark National Bank).

Drehman, Matthias, 2005, “A Market Based Macro Stress Test for the Corporate Credit Exposures of UK Banks,” available at http://www.bis.org/bcbs/events/rtf05Drehmann.pdf.

Elsinger, Helmut, Alfred Lehar, and Martin Summer, 2003, “Risk Assessment for Banking Systems” 14th Annual Utah Winter Finance Conference Paper; EFA 2003 Annual Conference Paper No. 437.

Eklund, Trond, Kai Larsen, and Eivind Berhardsen, 2003, “Model for analysing credit risk in the enterprise sector,” available at http://www.norges-bank.no/english/publications/economic_bulletin/2001-03/eklund-larsen.pdf.

Evjen, Snorre, Arild J Lund, Kjersti Haare Morka, Kjell B Nordal, and Ingvild Svendsen, 2005, “Monetary and Financial Stability in Norway: What Can We Learn from Macroeconomic Stress Tests?,” In: *Investigating the relationship between the*

financial and real economy, BIS Paper No. 22 (Basel: Bank for International Settlements), available at http://www.bis.org/publ/bppdf/bispap22u.pdf.

Gerlach, Stefan, Wensheng Peng, and Chang Shu, 2004, "Macroeconomic Conditions and Banking Performance in Hong Kong: A Panel Data Study," Hong Kong Monetary Authority Research Memorandum, April 2004, available at http://www.info.gov.hk/hkma/eng/research/RM_macro_and_banking.pdf.

Goodhart, Charles, 2006, "A Framework for Assessing Financial Stability?," *Journal of Banking and Finance*, forthcoming.

Hesse, Heiko, and Martin Čihák, 2007, "Cooperative Banks and Financial Stability," *IMF Working Paper* No. 07/02 (Washington: International Monetary Fund).

Hoggarth, Glenn, Andrew Logan, and Lea Zicchino, 2005, "Macro Stress Tests of UK Banks," *BIS Papers* No. 22, pp. 392–408.

Independent Evaluation Office (IEO) of the IMF, 2006, "Report on the Evaluation of the Financial Sector Assessment Program," available at www.imf.org.

International Association of Insurance Supervisors, 2003, "Stress Testing by Insurers: Guidance Paper," Guidance Paper No. 8, available at http://www.iaisweb.org/185stresstesting03.pdf.

International Monetary Fund, 2003, "Financial Soundness Indicators—Background Paper," May 14, 2003, available at http://www.imf.org/external/np/sta/fsi/eng/2003/051403bp.pdf.

——————, 2004, *Financial Soundness Indicators Compilation Guide*, available at http://www.imf.org/external/np/sta/fsi/eng/guide/index.htm

——————, 2005, "European Financial Integration, Stability and Supervision," in: "Euro Area Policies: Selected Issues," IMF Country Report No. 05/266, August, pp. 113–46, available at http://www.imf.org/external/np/sta/fsi/eng/guide/index.htm

——————, 2005b, "France: Financial Sector Assessment Program--Technical Notes--Stress Testing Methodology and Results; Integration into Global Financial Markets;and Public Intervention in Financial Markets--Obstacles to Monetary Transmission," at http://www.imf.org/external/pubs/ft/scr/2005/cr05185.pdf.

——————, 2006, "Spain: FSAP—Technical Note—Stress Testing Methodology and Results," IMF Country Report No. 06/216 (Washington: IMF), available at http://www.imf.org/external/pubs/ft/scr/2006/cr06216.pdf.

——————, and the World Bank, 2003, *Analytical Tools of the FSAP*, available at http://www.imf.org/external/np/fsap/2003/022403a.pdf.

———, 2005a, “Financial Sector Assessment Program—Review, Lessons, and Issues,” available at http://www.imf.org/External/np/fsap/2005/022205.htm.

———, 2005b, *Financial Sector Assessment: A Handbook*, available at www.imf.org/external/pubs/ft/fsa/eng/index.htm.

Jones, Matthew, Paul Hilbers, and Graham Slack, 2004, “Stress Testing Financial Systems: What to Do When the Governor Calls,” IMF Working Paper No. 04/127 (Washington: International Monetary Fund).

Kraft, Evan, and Tomislav Galac, 2007, “Deposit Interest Rates, Asset Risk and Bank Failure in Croatia,” *Journal of Financial Stability*, Volume 2, Issue 4, pp. 337–355.

Lucas, Robert E., 1976, “Econometric Policy Evaluation: A Critique,” Carnegie-Rochester Conference Series on Public Policy 1, pp. 19–46.

Peng, Wensheng, Kitty Lai, Frank Leung, and Chang Shu, 2003, “The Impact of Interest Rate Shocks on the Performance of the Banking Sector,” Hong Kong Monetary Authority Research Memorandum, May 2003, available at http://www.info.gov.hk/hkma/eng/research/RM07-2003.pdf.

Pesola, Jarmo, 2005, “Banking Fragility and Distress: An Econometric Study of Macroeconomic Determinants,” Bank of Finland Research Discussion Papers No. 13 (Helsinki: Bank of Finland). http://www.bof.fi/eng/6_julkaisut/6.1_SPn_julkaisut/6.1.5_Keskustelualoitteita/0513netti.pdf

Sahajwala, Ranjana, and Paul Van den Bergh, 2000, “Supervisory Risk Assessment and Early Warning Systems,” Basel Committee on Banking Supervision Working Paper No. 4, available at http://www.bis.org/publ/bcbs_wp4.pdf.

Segoviano Basurto, Miguel A., and Pablo Padilla, 2006, “Portfolio Credit Risk and Macroeconomic Shocks: Applications to Stress Testing Under Data-Restricted Environments,” IMF Working Paper No. 06/283 (Washington: International Monetary Fund).

Shu, Chang, 2002, “The Impact of Macroeconomic Environment on the Asset Quality of Hong Kong's Banking Sector,” Hong Kong Monetary Authority Research Memorandum, December 2002, available at http://www.info.gov.hk/hkma/eng/research/index.htm.

Virolainen, Kimmo, 2004, “Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland,” Bank of Finland Discussion Paper No. 18/2004.