Thoughts on the proper design of macro stress tests

Petr Jakubík and Gregory D Sutton[1]

1. Introduction

This paper provides an introduction to macro stress tests and argues that these exercises could be an important part of enhanced macroprudential policies.[2] Macro stress tests are executed by financial sector supervisors and central banks usually with (1) the aid of key financial institutions and (2) objectives that are different to those of stress tests run by financial firms for internal risk management purposes. For example, macro stress tests were undertaken during the recent international financial crisis with the aim of restoring confidence in financial systems (Bank of England (2008); Board of Governors of the Federal Reserve System (2009a, 2009b); Committee of European Banking Supervisors (2010)). Macro stress tests are also conducted during "good times" to search for potential sources of systemic risk. No matter when they are undertaken, macro stress tests usually estimate the losses that a group of financial institutions considered key to the proper functioning of a financial system, usually a group of large banks, could suffer under adverse macroeconomic developments or other shocks.

Macro stress tests could be an important part of enhanced macroprudential policies, because the recent financial crisis has shown the stress testing practices of a number of large banks to be seriously flawed. The difficulties some of them had dealing with the risks that emerged proves this. However, it was suggested even earlier that there were problems with financial firms' stress testing practices. In particular, the UK Financial Services Authority (2006) indicated that it was surprised by the mildness of the firm-wide stress tests undertaken by some financial firms, and the agency concluded that financial institutions might be underestimating the likelihood of severe events. This might be considered a logical consequence of short human memories and the prolonged period of favourable economic conditions prior to the crisis. Interestingly, the possibility that short memories can undermine successful risk management is consistent with the findings of Jiménez and Saurina (2006).

While "short memories" might explain why the stress tests run by a number of financial institutions were not severe enough, and thus left them too exposed, we believe that another flaw in stress testing practices also played an important role. This is the tendency for banks' firm-wide stress tests to focus on either the trading or loan book but not the combined market and credit risk losses associated with adverse shocks. Yet banks with sizable trading positions could suffer significant losses in both their trading and loan books following large macroeconomic shocks. This is because large swings in financial market prices and rates, such as steep increases in interest rates and exchange rate depreciations, have the potential to negatively impact the creditworthiness of households and firms, in addition to causing significant trading book losses. It is also because severe economic downturns that would be expected to lead to substantial loan-book losses are often associated with sharp declines in equity prices, sharp increases in credit spreads and other changes in financial markets.

---

[1] Petr Jakubík is an economist at the ECB. Gregory Sutton is a senior economist at the Financial Stability Institute of the BIS. The views expressed in this paper are those of the authors and do not necessarily represent the views of the ECB or the BIS.

[2] For discussions of macroprudential policies see, for example, Kashyap et al (2008), Bank of England (2009), Committee on the Global Financial System (2010) and Moreno (2011).

Therefore, stress tests of a single book have the potential to seriously underestimate the total losses a bank could suffer under adverse macroeconomic developments.

In our view, these two flaws in the stress testing practices of large banks represent a serious threat to financial stability. This is in part because the other main way that banks measure and manage their market and credit risks, the use of quantitative risk management models (QRMMs) such as value-at-risk, does not adequately measure risks associated with extreme events. The proper management of these so-called tail risks must therefore rely on effective stress testing. Moreover, QRMMs usually focus on either market or credit risk but not the combined losses across trading and loan books.[3] We therefore believe that macro stress tests that are based on sufficiently severe, yet plausible, macroeconomic shocks and that induce large banks to estimate (at least) the sum of their market and credit risk losses could improve the resilience of financial systems. Macro stress tests of this type could thus be a useful part of enhanced macroprudential policies. This is the motivation for the discussion of macro stress tests and sources of systemic risk that follows in Sections 2 and 3. Section 4 provides a brief conclusion.

2. General macro stress testing methodology

A macro stress test can be bottom up, top down or a combination of the two approaches. A bottom-up macro stress test is based on the results of individual stress tests conducted by the financial institutions themselves; however, all the institutions involved would rely on the same assumptions about future economic developments that are produced by the supervisor (or central bank), who also exercises some control over the many assumptions and bank internal models underpinning the exercise. In contrast, a top-down macro stress test is completely designed and performed by the supervisor and the same assumptions and models are applied to all institutions' balance sheet data.

These issues, and others, are discussed in greater detail in the large literature concerned with macro stress testing. Because macro stress tests have been performed for a number of years under the Financial Sector Assessment Programmes (FSAPs) of the IMF and World Bank, these institutions have published much of this literature (see, among others, Blaschke et al (2001), Jones et al (2004), IMF and World Bank (2005) and Čihák (2007)). Of course, financial sector supervisors and central banks also conduct macro stress tests outside FSAPs as part of their efforts to safeguard financial stability. The key underlying assumptions and results of macro stress tests, both conducted within and outside FSAPs, are sometimes published in financial stability reviews.[4] As the literature on macro stress testing makes clear, insurance companies and other non-bank financial institutions are sometimes included in the group of firms covered by a macro stress test. However, because of the important role of banks in the provision of credit, and also their reliance on short-term funding, macro stress tests usually focus on a country's banking sector. Perhaps not surprisingly, therefore, an analysis of credit risk is usually an important part of a macro stress test. Because macroeconomic risk is arguably the main common source of loss for many credit exposures, macro stress tests almost always estimate the impact an economic downturn (or slowdown) would have on banks' credit losses. Other risks may also be covered by the exercise, as we discuss in Section 3.

---

[3] An exception is the systemic risk model employed by the Central Bank of the Republic of Austria (see Boss et al (2006)).

[4] See, for example, Boss (2002), Hoggarth and Whitley (2003), De Bandt and Oung (2004), Bunn et al (2005), European Central Bank (2009) and the Central Bank of Norway (2010).

Sensitivity tests and scenario analysis

There are two main ways to form the adverse shocks that underpin stress tests.[5] One approach assumes that only a single risk factor undergoes a significant change. These are known as sensitivity tests. An advantage of these tests is that, because only a single variable is shocked, they may be relatively easy to implement. Unfortunately, however, sensitivity tests may lack plausibility, because in a stress event it is unlikely that only a single key variable will be significantly affected. Nevertheless, many macro stress tests still rely on a single-variable shock as their starting point.

A more plausible approach to stress testing is a scenario analysis that examines the impact of changes in a number of key variables. Because it is a more plausible approach, and more likely to lead to an accurate estimate of the sum of credit and market risk losses under adverse developments, it is our preferred starting point for a macro stress test. Of course, it is more difficult to specify how a number of variables would move together during a stress event. The use of macroeconomic models can help address this challenge, because they can restrict the co-movements of variables to be consistent with economic theory.

The variables typically shocked in macro stress tests, in either sensitivity tests or scenarios, are interest and exchange rates, measures of inflation and unemployment, GDP and property prices. A central bank's official macroeconomic forecast, obtained from its macroeconomic model, usually serves as the starting point for deciding on shock sizes. Shocks should be, while plausible, also large. This is because large shocks are more likely to pose a significant threat to financial stability and also because they may not be adequately reflected in firms' internal risk management practices.

Assessing the impact of a shock

Whether a single- or multi-variable shock underpins a macro stress test, it is important to specify an appropriate time horizon over which the effects of the disturbance will be traced. An appropriate time horizon will balance competing forces. On the one hand, it probably takes a relatively long time for most of the credit losses associated with an adverse shock, such as a significant decline in domestic economic activity, to be realised. Given that most macro stress tests aim to include estimates of losses from credit exposures, this argues for a relatively long time horizon. On the other hand, a shorter time horizon makes it less important to model changes in financial institutions' portfolios. In practice, the time horizon of a macro stress test is usually between one and three years.

Another issue is the metric used to evaluate whether a financial institution or system would be able to absorb a particular measure of loss, be it expected loss or the maximum loss with some probability. The standard metric is to assess the measure of loss relative to capital or assets, taking into consideration estimates of current and future net income of financial institutions. Forecasts of net income are commonly based only on past average income. While it would be preferable to model how a large number of components of income would evolve under the stress scenario, this is no easy task. This is because the incomes of financial institutions depend in complicated ways on a large number of factors, including the extent to which income sources are diversified. This is unfortunate because, in the event of losses, net income protects capital.

[5] The same two approaches underpin macro stress tests and the stress tests financial institutions run for internal risk management purposes.

3. Sources of systemic risk

Macro stress tests are useful for shedding light on potential sources of systemic risk. Systemic risk can be appreciable when important financial institutions have large, common exposures to macroeconomic developments, financial market prices or real estate prices. "Short memories" could play an important role in the build-up of large, common exposures (especially after a prolonged period of favourable economic conditions). Systemic risk can also arise from self-reinforcing feedback loops. In the remainder of this section we discuss a number of these risks in greater detail and, for the less difficult to quantify, how a macro stress test might attempt to measure them.

Interest and exchange rate risk

Financial institutions are likely to have common exposures to interest rate, and perhaps also exchange rate, risk. Most trading portfolios of financial institutions would probably suffer from unexpected increases in interest rates. The financial impact of parallel upward (and downward) shifts of government yield curves (base rates) can be approximated, following Macaulay, as the product of the interest rate change, duration and original value of securities. The financial impact of sharp rises in credit spreads should probably also be investigated, because large increases in spreads are more likely than large declines.[6] An estimate of exchange rate risk can be obtained by multiplying the net open FX positions of financial institutions, both on- and off-balance sheet, by assumed changes in key exchange rates. There is also an element of interest rate and exchange rate risk that shows up as credit risk for financial institutions, as will be discussed below.[7]

Real estate price risk

Real estate price movements can also be a source of systemic risk, in part because of the importance of real estate as collateral for loans from banks and other creditors. While financial sector losses are usually associated with declines in the prices of commercial and/or residential real estate, systemic risk can be increasing during a prolonged period of rising real estate prices. This is especially the case if real estate markets get caught up in a bubble and prices rise significantly more than justified by economic fundamentals. In this case the eventual bursting of the bubble would likely be associated with marked falls in the prices of real estate, an elevated incidence of default by owners of residential and commercial properties and perhaps also reduced recoveries for financial institutions and other creditors in the case of default.

Rising systemic risk associated with an inflating house price bubble was arguably the situation in the United States in the years just prior to the recent international financial crisis. As the bubble began to deflate after 2006, an unexpectedly large number of homeowners found themselves in a position of negative equity and chose to default. Given the high loan-to-value ratios of many subprime home loans, the US subprime sector was where problems first emerged. However, as housing prices continued to fall in the United States, defaults also rose significantly for homes purchased with traditional mortgages.

---

[6] In the case of banks, there is also interest rate risk in their loan books which shows up in volatility of net interest income.

[7] Exchange rate risk can also emerge as funding liquidity risk, a risk that we do not cover. For discussions of approaches to incorporate funding liquidity risk into macro stress tests, see Čihák (2007) and Aikman et al (2009).

Credit risk in the loan book

Another key risk from a systemic perspective is the credit risk in large banks' loan (or banking) books. Significant declines in real estate prices can be an important cause of losses in loan books; however, the most important influence is often the state of the macro economy, reflected in large part by the growth of domestic national income and perhaps also national income abroad through its effect on exports.[8] All else being equal, more rapid income growth both at home and abroad would arguably make it easier for corporate and household borrowers to service their debts. Changes in interest and exchange rates can also influence the creditworthiness of banks' counterparties. In the case of households and non-financial firms, increases in interest rates might be expected to decrease their creditworthiness, especially if a significant amount of their debt is in floating rate agreements.[9] It is less clear what moves in exchange rates might do most damage to a bank's loan book. In the absence of complete currency hedging, exporters with negligible foreign currency debts would likely become greater credit risks when the domestic currency appreciates. However, when companies and households have significant foreign currency debts, but insignificant foreign currency income, a depreciation of the domestic currency can seriously damage their creditworthiness.[10]

One way of measuring credit risk in loan books is to project losses under an adverse macroeconomic scenario. In bottom-up macro stress tests, calculations are based on banks' own internal models using common scenarios for all of them. In top-down macro stress tests, regression methods applied to aggregate data can be used to project how the particular measure of credit losses, for the economy as a whole or a group of banks, would evolve under assumed paths for key macroeconomic variables such as GDP, real estate prices, interest rates and exchange rates.[11] Often the ratio of non-performing loans (NPL) to total loans is the measure of credit losses in top-down tests; however, this indicator can significantly underestimate credit risk in the case of rapid credit growth, because it usually takes time for bad loans to reveal themselves.

Another approach to measuring credit risk in loan books is macroeconomic credit risk modelling. This approach, pioneered by Wilson (1997), relates credit risk to the "health" of the domestic economy as revealed by a number of macroeconomic variables. In the default mode application of the model, probabilities of default (PDs) for different classes of borrowers are related to macroeconomic variables, and the time series dynamics of the macroeconomic variables are also modelled. The approach is thus capable of predicting how PDs will evolve in the future under alternative macroeconomic scenarios, given initial macroeconomic conditions. These starting conditions could be the actual values of macroeconomic variables or those that would be associated with a particular stress scenario. Using further simulations, the conditional loss distribution associated with a bank's loan book can be derived.

In the case of corporate exposures, it may be possible to perform a more refined analysis of expected losses under a stress scenario with the use of statistical credit rating models.

---

[8] The state of the macro economy is of course an important determinant of real estate prices, in addition to being a key determinant of credit risk more broadly.

[9] Even if this is not the case, the fact that interest rate increases mean that new borrowing would have to take place at higher rates reduces creditworthiness, because it reduces an entity's ability to survive adverse shocks.

[10] Of course, importers without significant foreign currency debts could also see their creditworthiness decline in response to a depreciation of the domestic currency.

[11] Pesola (2001), Shu (2002), Pain (2003) and Jakubík and Schmieder (2008) provide examples of statistical models relating measures of loan-book losses to macroeconomic variables.

These models associate with each exposure a score which is then related to the exposure's PD. An expected loss for the exposure can be calculated as a product of the PD, the exposure at default (EAD) and the loss-given-default (LGD). Statistical credit rating models include obligor-specific data when generating scores. When these models also incorporate macroeconomic variables, they can be used to compute "stressed" scores and PDs for obligors that would be associated with adverse macroeconomic developments.[12] These stress losses for individual exposures can be aggregated to obtain total stress losses for the corporate loan book.

Credit risk parameters

Short time series and structural breaks can prevent the use of some of the methods for evaluating credit risk discussed above. In this case, expected credit losses for an entire portfolio can be calculated as the product of the portfolio-average PD and LGD and the total portfolio EAD. Perhaps the biggest challenge with this approach is to determine appropriate LGDs, and this parameter is very often determined mainly by expert judgment. Real estate prices can often inform this judgment, because of real estate's important role as collateral.

EAD can be expressed as the difference between outstanding loans and NPL. Expected NPL depend on the inflow to NPL (determined by PD estimates), outflows (as written-off or sell-out of existing NPL) and the current stock of NPL. Formally,

$$NPL_{t+1} = NPL_t + PD_t \cdot (Loans_t - NPL_t) - r \cdot NPL_t \quad (1)$$

where *r* represents the average write-off (or sell-out) rate of existing NPL.[13]

Expert judgment can also be used to link the PD of a portfolio to a single macroeconomic variable; alternatively, when a regression model for the growth of NPL is estimated, the PD for the portfolio can be easily approximated. This can be very helpful in the case of emerging economies where aggregate data on defaults are very often not available, but aggregate data on NPL are available. To see this, assume that the stock of NPL is relatively small compared to the stock of loans, so that equation (1) can be approximated by:

$$NPL_{t+1} \approx NPL_t + PD_t \cdot Loans_t - r \cdot NPL_t \quad (2)$$

Rearranging (2) gives:

$$PD_t \approx \left(\frac{\Delta NPL_{t+1}}{NPL_t} + r\right) \cdot \frac{NPL_t}{Loans_t} \quad (3)$$

The approximation (3) demonstrates that the average PD depends on the growth rate of NPL, the average write-off rate and the initial level of NPL relative to outstanding loans.

Self-reinforcing feedback loops

Self-reinforcing feedback loops can also be a source of systemic risk. One such loop can arise when banks and other financial institutions experience large, unexpected losses. It would be natural to expect that lending standards might be tightened in response, in part to rebuild capital buffers but also as a consequence of the increase in perceived risk, and perhaps also elevated risk aversion, that would likely follow such losses. This could lead to a

[12] See, for example, Vallés (2006).

[13] This parameter is often set to its average value over a previous period, usually several years, although this is recognised to be only a rough approximation to actual bank behaviour.

further deterioration in the real economy, additional financial sector losses and further cuts in credit availability.[14] A self-reinforcing feedback loop can also result from asset fire sales.[15]

Although potentially important, self-reinforcing feedback loops are typically not taken into formal consideration in macro stress testing exercises, in part because such nonlinearities are difficult to model. Concern about them, and the systemic risk they potentially generate, can nevertheless be a motivation for macro stress tests. As noted by Bernanke (2010), one of the objectives of a macro stress test recently carried out in the United States (SCAP) was to ensure that large US banks would continue to lend to creditworthy households and firms even if economic conditions turned out worse than expected. To the extent that the SCAP was credible in this respect, expectations were more likely concentrated on a more favourable future macroeconomic trajectory, with lower financial sector losses, increasing confidence in the health of the US financial system.

4. Conclusion

This paper provides an introduction to macro stress tests and some of the risks that they attempt to assess. It also argues that these exercises could improve the resilience of financial systems because, in our view, banks' stress testing practices are seriously flawed. One reason is that the assumed developments underpinning stress tests appear not to be severe enough. Another is the tendency for banks' stress tests to focus on either the trading or loan book but not the combined market and credit risk losses likely to accompany adverse macroeconomic shocks. We believe that macro stress tests that have severe, yet plausible, multi-variable macroeconomic shocks as their starting point and estimate (at least) market and credit risk losses could therefore support the preservation of financial stability. Macro stress tests of this type could thus be a useful part of enhanced macroprudential policies.

References

Aikman, D, P Alessandri, B Eklund, P Gai, S Kapadia, E Martin, N Mora, G Sterne and M Willison (2009): "Funding liquidity risk in a quantitative model of systemic stability", *Bank of England Working Paper*, no 372.

Bank of England (2008): *Financial Stability Report*, October.

——— (2009): "The role of macroprudential policy", Discussion Paper, November.

Bernanke, B (2010): "The supervisory capital assessment program – one year later", Speech at the Federal Reserve Bank of Chicago 46th Annual Conference on Bank Structure and Competition, May.

Blaschke, W, M Jones, G Majnoni and M Peria (2001): "Stress testing of financial systems: an overview of issues, methodologies, and FSAP experiences", *IMF Working Paper*, no 01/88.

[14] As discussed by Borio, Furfine and Lowe (2001), risk-sensitive regulatory capital requirements and the dependence of collateral values on the state of the macroeconomy may also play a role in this self-reinforcing feedback loop.

[15] For a discussion of an approach to incorporate asset market liquidity and fire-sale effects into macro stress tests, see Aikman et al (2009).

Board of Governors of the Federal Reserve System (2009a): “The supervisory capital assessment program: design and implementation”.

——— (2009b): “The supervisory capital assessment program: overview of results”.

Borio, C, C Furfine and P Lowe (2001): “Procyclicality of the financial system and financial stability: issues and policy options”, *BIS Papers*, no 1.

Boss, M (2002): “A macroeconomic credit risk model for stress testing the Austrian credit portfolio”, *Financial Stability Report* 4, Central Bank of the Republic of Austria, pp 64–82.

Boss, M, G Krenn, C Puhr and M Summer (2006): “Systemic risk monitor: a model for systemic risk analysis and stress testing of banking systems”, *Financial Stability Report* 11, Central Bank of the Republic of Austria, pp 83–95.

Bunn, P, A Cunningham and M Drehmann (2005): “Stress testing as a tool for assessing systemic risks”, *Bank of England Financial Stability Review*, June, pp 116–26.

Central Bank of Norway (2010): *Financial Stability*, Report, November.

Čihák, M (2007): “Introduction to applied stress testing”, *IMF Working Paper*, no 07/59.

Committee of European Banking Supervisors (2010): “Aggregate outcome of the 2010 EU-wide stress testing exercise coordinated by CEBS in cooperation with the ECB”.

Committee on the Global Financial System (2010): “Macroprudential instruments and frameworks: a stocktaking of issues and experiences”, *CGFS Papers*, no 38.

De Bandt, O and V Oung (2004): “Assessment of “stress tests” conducted on the French banking system”, *Bank of France Financial Stability Review*, November.

European Central Bank (2009): *Financial Stability Review*, December.

Financial Services Authority (2006): “Stress testing thematic review”, *Dear CEO Letter*, October.

Hoggarth, G and J Whitley (2003): “Assessing the strength of UK banks through macroeconomic stress tests”, *Bank of England Financial Stability Review*, June, pp 91–103.

IMF and World Bank (2005): *Financial Sector Assessment: A Handbook*, International Monetary Fund and the World Bank.

Jakubík, P and C Schmieder (2008): “Stress testing credit risk: comparison of the Czech Republic and Germany”, Financial Stability Institute Award 2008 Winning Paper.

Jiménez, G and J Saurina (2006): “Credit cycles, credit risk, and prudential regulation”, *International Journal of Central Banking*, June.

Jones, M, P Hilbers and G Slack (2004): “Stress testing financial systems: what to do when the governor calls”, *IMF Working Paper*, no 04/127.

Kashyap, A, R Rajan and J Stein (2008): “Rethinking capital regulation”, in *Maintaining Stability in a Changing Financial System*, Federal Reserve Bank of Kansas City, pp 431–71.

Moreno, R (2011): “Policymaking from a “macroprudential” perspective in emerging market economies”, *BIS Working Papers*, no 336.

Pain, D (2003): “The provisioning experience of the major UK banks: a small panel investigation”, *Bank of England Working Paper*, no 177.

Pesola, J (2001): “The role of macroeconomic shocks in banking crises”, *Bank of Finland Discussion Papers*, no 6/2001.

Shu, C (2002): “The impact of macroeconomic environment on the asset quality of Hong Kong’s banking sector”, Hong Kong Monetary Authority Research Memorandum, December.

Vallés, V (2006): “Stability of a “through-the-cycle” rating system during a financial crisis”, Financial Stability Institute Award 2006 Winning Paper.

Wilson, T (1997): “Portfolio credit risk (I)”, *Risk*, 10(9), September, pp 111–7.