Stress testing credit risk: modelling issues

Stijn Ferrari
Patrick Van Roy
Cristina Vespro

Introduction

National supervisory authorities – partially under the impetus of the IMF-World Bank Financial Sector Assessment Program (FSAP) – have been conducting stress testing exercises in their respective jurisdictions for almost a decade. However, the recent financial crisis has highlighted stress testing as an increasingly important prudential tool for assessing the banking system's resilience to possible adverse economic developments and shocks concerning a variety of risks, such as credit, market and liquidity risks. To complement national analyses, in 2009 the Committee of European Banking Supervisors (CEBS) – with the support of the European Central Bank (ECB) – launched an annual coordinated assessment of European banks using common scenarios and methodologies. Following the example of the Supervisory Capital Assessment Program (SCAP) conducted in the US in 2009, the general methodology and both aggregate and bank-specific results of the 2010 CEBS EU-wide stress test were publicly disclosed, with the aim of increasing market confidence and reducing uncertainties surrounding major European banks' fragilities. In addition, the regulation establishing the European Banking Authority (EBA) specifies that the EBA will initiate and coordinate regular, EU-wide stress tests to assess the resilience of financial institutions. Hence, EU-wide coordination and a high level of transparency are characteristics of the EBA's 2011 continuation of the EU-wide stress testing exercise.

Stress tests provide estimates of the impact of stressed macroeconomic scenarios on banks' health. The macroeconomic stress is likely to have an impact on the components of the profit and loss account (P&L), which through retained earnings impact the solvency position of banks' balance sheets. Despite ongoing efforts to incorporate assessments of multiple bank risks in stress testing frameworks, credit risk remains the main type of bank risk that is assessed in bank supervisory authorities' stress testing exercises. As a consequence, most stress testing exercises concentrate on the impairments on banks' loan portfolios (as opposed, for example, to exposures held for trading purposes).

In considering the impact of a stressed macroeconomic scenario on banks' loan impairments, the critical question that must be addressed by those conducting the stress test is how to measure the impact of the stressed macroeconomic variables on credit risk associated with the loan portfolio. Ideally, one would want to translate shocks to macroeconomic variables into an increase in expected losses on the portfolio. However, there are no unanimously accepted techniques for making such estimates. As a result, a number of modelling decisions must be made in the course of a stress test. These decisions thus obviously introduce model risk into the outcome of stress tests (i.e., into the estimated impact on banks).

This article discusses a number of questions related to model risk in stress testing that have not been widely discussed but that nevertheless need to be addressed and understood by practitioners of stress tests. In particular, more needs to be understood about the relative implications of differing modelling choices in the stress testing process. These modelling choices relate to the following basic questions: (i) How to link credit risk to the macroeconomic environment?, and (ii) Which variable(s)
to use for measuring credit risk in the portfolio? While these questions seem elementary, practitioners’ modelling choices and practical issues such as data availability have resulted in different approaches being taken, both in practice and in the literature on stress testing (see the table in the Appendix for an overview of and references to the different approaches). In our opinion, insufficient attention has been devoted to how the choices on the above two questions may affect the practical implementation and the outcome of stress tests.

The remainder of the article is organized as follows. Section 1 presents the general modelling framework used in credit stress testing and discusses the impact of modelling choices on the implementation of stress testing as well as the risk of model inconsistency. In Section 2 we provide an overview of the credit risk variables that can be used for stress testing and the issue of robustness of stress testing outcomes across different credit risk variables. Section 3 offers some concluding remarks.

1. The modelling framework

In this first Section we discuss the first of our two basic questions, namely how to link credit risk to the macroeconomic environment. As a general background, we provide a brief overview of the different components of the credit stress testing process. A discussion of the credit risk part of the 2011 EBA stress testing exercise in this context is provided in Box 1. Next, we discuss how the modelling framework may affect the practical implementation of stress testing and warn against the potential for model inconsistency that may arise in this context.

1.1 The stress testing process

Credit stress tests assess the impact of a stressed macroeconomic scenario on the quality of banks’ credit exposures. This approach essentially consists of three steps: (i) forecast values of macroeconomic variables under a given pre-specified (stressed) scenario over a given horizon, (ii) estimate the impact of the stressed macroeconomic variables on the banks’ credit risk parameters (typically probability of default (PD) and loss given default (LGD)) over a given horizon, and (iii) use these stressed credit risk parameters to evaluate the impact of the stress scenario on the banks’ P&L (and solvency). This typical stress testing process is summarized in Chart 1.

The practical implementation of steps (i) and (ii) generally involves two different modelling stages. First, for forecasting the behaviour of macroeconomic variables (e.g., GDP growth and the long-term interest rate) under a pre-specified stress scenario over a given horizon, a macroeconometric model is typically used. In general,

(1) Descriptions of the credit stress testing process and stress testing processes in general can be found in for example in Blaschke et al. (2001), Jones, Hilbers and Slack (2004), Čihák (2007) and Foglia (2009).

---

**Chart 1**  THE TYPICAL CREDIT STRESS TESTING PROCESS

- **Stress scenario** (e.g., oil price shock, aggregate demand shock)
  - Macroeconometric model (structural or reduced-form)
    - determines stressed value of the macroeconomic variables (e.g., GDP growth, the long-term interest rate)
  - Credit risk model
    - determines stressed value of the credit risk variable
  - Bank’s P&L and solvency
adverse shocks to one or more macroeconomic variables are entered into the model, and the equations in the model determine how these and other macroeconomic variables behave over the stress testing horizon as a consequence of the shocks. A second modelling component is required to estimate the impact of some stressed macroeconomic variables of interest(1) on the banks’ credit risk parameters over the stress testing horizon. This component, the credit risk model, essentially consists of one or more equations linking the banks’ credit risk parameters to the macroeconomic variables. The stressed macroeconomic variables obtained from the macroeconometric model are entered into this equation to obtain stressed values of the credit risk parameters. Finally, the stressed values of the credit risk parameters are applied to the banks’ P&L to obtain the estimated impact on their solvency position.

A distinction may be made between “top-down” and “bottom-up” stress tests. In their purest forms, the top-down stress test is one in which authorities use data at their disposal to conduct all the steps of the stress test, including estimating the ultimate impact on P&L and solvency positions. In a pure bottom-up stress test, banks estimate the stressed credit risk variables and the ultimate impact on their P&L and solvency. Sometimes, a stress test is referred to as bottom-up if banks perform the final step, even though authorities may have supplied to the banks the stressed values of the macroeconomic variables and the stressed values of credit risk variables.

As macroeconometric models typically do not include a detailed financial sector and credit-risk related variables, the two modelling components outlined above are usually considered as two separate entries or modules in the stress testing framework. In this “modular approach”, the credit risk model is often referred to as the “satellite (or auxiliary) model”. When in contrast the link between credit risk and macroeconomic developments is integrated into or jointly estimated with the macroeconometric model, we refer to the stress testing framework as following an “integrated approach” (as indicated by the blue dashed Box in Chart 1).

In the following subsections we discuss some of the implications of opting for a modular approach as compared to an integrated approach. First, we explain how the practical implementation of stress testing may either result in a point estimate of stressed credit risk or in an entire distribution of values for a stressed credit risk variable, in which case the focus can be on the tails of the distribution (i.e., the extreme observations in the distribution). The choice of either a modular or an integrated approach affects the degree to which relationships that are not captured by the equations of the models can be taken into account in the generation of the credit risk distribution. Second, we point at potential sources of model inconsistency that may arise especially in the modular approach and threaten the internal consistency of the stress testing modelling framework.

---

(1) The macroeconomic variables of interest to which the credit risk parameters are linked are selected on the basis of economic theory and statistical significance.

---

Box 1 – The credit risk part of the 2011 EBA stress testing exercise

The 2011 EBA stress testing exercise embodies an internationally coordinated assessment, using common scenarios and methodologies, of European banks’ resilience to possible adverse economic developments and their ability to absorb possible shocks on credit, market and funding risks. In a centralized approach, coordination takes place between the EBA (with the support of the ECB) and national central banks/supervisory authorities, and between national central banks/Supervisory authorities and the banks included in the stress testing exercise. The overall objective of the exercise is to provide policy information for assessing the resilience of the EU banking system.

The credit risk part of the 2011 EBA stress testing exercise is very much in line with the typical credit stress testing process described in the main text. In a first step, forecasts for the macroeconomic variables – reflecting a benchmark and an adverse scenario over a two-year horizon – are obtained from macroeconometric models by the European Commission and the ECB. In a second step, the credit risk parameters (PD and LGD) of different bank portfolios (financial institutions, sovereign, corporate, consumer credit and retail real estate) are evaluated under the two scenarios via a credit risk model linking these two parameters to a set of macroeconomic variables. Finally, the impact of the macroeconomic scenarios on the banks’ solvency positions is assessed.
Once the necessary links between the macroeconomic environment and credit risk have been established via the modelling framework, the effect on credit risk of a given macroeconomic scenario can be estimated. As mentioned above, one or more shocks are typically entered into the macroeconometric model in order to generate a stressed path of some macroeconomic variables of interest. These stressed macroeconomic variables obtained from the macroeconometric model are then entered into the credit risk model to obtain stressed values of the credit risk parameters. Several approaches have been followed in implementing this in practice and in the credit stress testing literature. One can distinguish between the “deterministic approach” and the “stochastic (or probabilistic) approach”. The choice of approach determines whether one obtains a single estimate of the realisation of a stressed credit risk variable or a distribution of potential realisations of the variable.

We illustrate the difference between the two approaches on the basis of Chart 2. The Chart depicts the practical implementation of the stress testing process using a simplified macroeconometric model that determines the behaviour of two macroeconomic variables (GDP growth and the long-term interest rate) and a credit risk model that links the credit risk variable to the macroeconomic variables. To keep our example simple, suppose that the underlying model equations are the following:

\[
\begin{align*}
\text{GDP} &= f_{\text{GDP}}(\text{GDP}, i, \varepsilon_{\text{GDP}}) \\
i &= f_i(\text{GDP}, i, \varepsilon_i) \\
C &= f_C(\text{GDP}, i, \eta)
\end{align*}
\]

where:

- \(\text{GDP}\) is GDP growth,
- \(i\) is the long-term interest rate,
- \(C\) is the credit risk variable (e.g., bankruptcy rates, loan loss provisions or non-performing loans\(^{(1)}\)),
- \(\varepsilon_{\text{GDP}}\) and \(\varepsilon_i\) represent error terms in the estimating regressions for \(\text{GDP}\) and \(i\), respectively,
- \(\eta\) represents the error term in the estimating regression for the credit risk variable \(C\).

Thus, in this simple model, GDP growth is modelled by using its own lagged values, and the current value of the long-term interest rate as explanatory variables. Similarly, the long-term interest rate is modelled using its own lagged values and the current value of GDP growth as explanatory variables. Finally, in the credit risk model, the credit risk variable is modelled with the current values of GDP growth and the long-term interest rate as explanatory variables. The error terms in each equation reflect the variation in the value of the dependent variable not explained by the explanatory variables. The model, finally, entails making an assumption about the distribution of the error terms (for example, the normal distribution).

In the deterministic approach (as illustrated by the blue-shaded boxes in Chart 2), a stressed realisation of each of the error terms \(\varepsilon_{\text{GDP}}\) and \(\varepsilon_i\) is drawn from their assumed distribution. The values of these realisations \((\varepsilon_{\text{GDP}}^1, \varepsilon_i^1)\) are then used in the estimated equations of the macroeconometric model to obtain forecast values of GDP growth \((\text{GDP}^1)\) and the long-term interest rate \((i^1)\). These values are then inserted in the estimated equation for the credit risk variable in order to obtain an estimate of the stressed value of the credit risk variable \((C^1)\). Thus, a single estimate of the stressed credit risk variable is obtained based on taking the fitted values of the credit risk equation\(^{(2)}\), where the macroeconomic variables are evaluated at their simulated stressed values and where the error term \(\eta\) is set equal to zero.

---

\(^{(1)}\) See Section 2 for a discussion of potential credit risk variables.

\(^{(2)}\) In practice, as the stress testing horizon contains multiple periods, one particular path of the stressed credit risk variable is based on one particular path of the stressed macroeconomic variables.
In contrast, the stochastic approach is characterised by repeating the above process many times, thereby generating a distribution of the credit risk variable, i.e., many different realisations rather than a single point estimate. This credit risk distribution can be obtained in several ways. One of the possible approaches is illustrated in Chart 2. Rather than taking one shock for \( \epsilon_{GDP} \) and \( \epsilon_i \) (i.e., the realisations \( \epsilon_{GDP}^1 \) and \( \epsilon_i^1 \)) in the macroeconometric model to obtain single values of the stressed macroeconomic forecasts \( GDP^1 \) and \( i^1 \), and consequently a single value for the stressed credit risk variable \( C^1 \), one can repeat this process many (e.g., 10 000) times. As indicated by the blue dashed boxes and arrows, each set of macroeconomic shocks results in a predicted value of the stressed credit risk variable. The distribution implied by the 10 000 stressed credit risk variable forecasts \( C^1, ..., C^{10 000} \) may then be used to obtain a distribution of expected credit losses and allows focusing on the extreme observations within this distribution (the tail of the distribution). However, potential correlations between the error terms of the macroeconometric equations (\( \epsilon_{GDP} \) and \( \epsilon_i \)) and the credit risk equation (\( \eta \)) are not accounted for when generating the credit risk distributions.

An alternative approach to obtain a distribution rather than a point estimate of the credit risk variable is to insert into the credit risk model the stressed forecasts for GDP growth and the long-term interest rate \( (GDP^1 \text{ and } i^1) \) based on one particular set of shocks \( \epsilon_{GDP}^1 \) and \( \epsilon_i^1 \) in the macroeconometric model, as is done in the deterministic approach, and then augment this step by taking many (e.g., 10 000) draws for the error term \( \eta \) of the credit risk model (rather than setting this term equal to zero, as was the case for the previous methodology). Given the macroeconomic stressed forecasts \( GDP^1 \) and \( i^1 \), each draw of \( \eta \) results in a stressed value for the credit risk variable. Again, potential correlations between the error terms of the macroeconometric equations (\( \epsilon_{GDP} \) and \( \epsilon_i \)) and the credit risk equation (\( \eta \)) are not accounted for when generating the credit risk distributions.

A combination of both approaches described above is a straightforward extension (1), in that multiple draws can be taken for the error terms \( \epsilon_{GDP} \) and \( \epsilon_i \) in the macroeconometric model, while at the same time taking many draws for the error term \( \eta \) of the credit risk model. Ideally, correlations between the error terms of the macroeconometric equations and the credit risk equation (i.e., between \( \epsilon_{GDP} \) and \( \epsilon_i \) on the one hand and \( \eta \) on the other) are taken into account in this approach. This would allow for the

---

(1) A distribution of the credit risk variable may also be obtained by applying bootstrapping methods. However, the purpose of this type of exercise is rather to evaluate the robustness and consistency of the underlying model specification.
possibility that the macroeconomic shocks affect the credit risk variable not only through the observable variables of the modelled equations, but also through their correlation with the error term in the credit risk model. For example, when both the macroeconomic variables and the credit risk variable depend on some factors that are not taken into account in the models, the effect of these common factors may not be captured by the equations of the different models, but rather be embedded in the relationships between the error terms. In our simple example with three variables (GDP growth, the long-term interest rate and the credit risk variable), such a common factor may be oil prices, for instance. Oil prices may affect not only GDP growth and the long-term interest rate, but also the debt servicing capacity of the banks’ credit counterparts (and therefore the credit risk variable). The effects of oil prices are not reflected in the equations of our simplified modelling framework, but may be taken into account when allowing for correlations between the error terms of the macroeconomic model and the credit risk model. In principle, an integrated approach is the only one which enables these correlations to be taken into account when generating values for the error terms of the macroeconomic equations and the credit risk equation. In a modular approach, the macroeconomic model and the credit risk model are treated as separate components or modules in the stress testing framework. The importance of taking account of these correlations in existing credit stress testing models is an issue which has hardly been investigated as yet.

1.3 Risk of model inconsistency

The risk of model inconsistency arises when separate model components in the stress testing framework, in casu the macroeconomic model and the credit risk model, contain the same or related variables, i.e., when equations of the same variables are estimated both in the macroeconomic model and in the credit risk model. In this case, discrepancies may arise between the forecasts implied by the credit risk model and those determined on the basis of the macroeconomic model.

We illustrate this point by building an example on the basis of the simplified model discussed in the previous subsection. The left-hand panel of Table 1 repeats the equations of this model, which consists of two macroeconomic equations (explaining GDP growth and the long-term interest rate) and one credit risk equation (explaining the credit risk variable using the macroeconomic variables).

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>ILLUSTRATION OF POTENTIAL MODEL INCONSISTENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplified model</td>
<td>Feedback model</td>
</tr>
<tr>
<td>macro-economic model</td>
<td>macro-economic model</td>
</tr>
<tr>
<td>( GDP = f_{\text{DEP}}(GDP, i, \varepsilon_{\text{DEP}}) )</td>
<td>( GDP = f_{\text{DEP}}(GDP, i, \varepsilon_{\text{DEP}}) )</td>
</tr>
<tr>
<td>( i = f_{i}(GDP, \varepsilon_{i}) )</td>
<td>( i = f_{i}(GDP, \varepsilon_{i}) )</td>
</tr>
<tr>
<td>credit risk model</td>
<td>credit risk model</td>
</tr>
<tr>
<td>( C = f_{C}(GDP, i, \eta) )</td>
<td>( C = f_{C}(GDP, i, \eta) )</td>
</tr>
<tr>
<td>( GDP = g_{\text{DEP}}(GDP, i, \varepsilon_{\text{DEP}}) )</td>
<td>( GDP = g_{\text{DEP}}(GDP, i, \varepsilon_{\text{DEP}}) )</td>
</tr>
<tr>
<td>( i = g_{i}(GDP, \varepsilon_{i}) )</td>
<td>( i = g_{i}(GDP, \varepsilon_{i}) )</td>
</tr>
</tbody>
</table>

Suppose now that the practitioner conducting the stress test (more specifically, the person running the credit risk model), has to take the output of the macroeconomic model (i.e., the first two equations of our simplified model) as given. For example, the practitioner may be a member of a supervisory authority or of the financial stability department of a central bank, and the stress test will make use of a macroeconomic scenario generated by the macroeconomic forecasting unit of the central bank or by some external provider.

Suppose in addition that the practitioner’s model for generating the stressed values of the credit risk variable includes extra equations, for example to capture feedback effects from the credit risk variable to GDP growth and the long-term interest rate, as illustrated in the right-hand panel of Table 1.\(^{(1)}\) There is now a risk of model inconsistency, since the additional macroeconomic equations in the credit risk model may differ from those in the macroeconomic model, which could potentially result in predicted stressed values for the macroeconomic variables different from those implied by the macroeconomic model.

Another example of potential model inconsistency arises when the macroeconomic model includes an equation with a credit risk variable, perhaps measured at a very aggregate level (such as economy-wide default rates of firms), while the practitioner’s credit risk equation is estimated at some lower level of aggregation (e.g., sectoral), and may also include different explanatory variables. The presence of two separate credit risk equations, one in the macroeconomic model and the other in the credit risk model, again gives rise to potential model inconsistency.\(^{(2)}\)

---

\(^{(1)}\) Stress testing applications that separately model macroeconomic variables in addition to using a macroeconomic model for stressed predicted values include van den End, Hoeberichts and Tabbae (2006) and Åsberg and Shahnazarian (2008).

\(^{(2)}\) In Andersen et al. (2008), for example, firm and household default rates are modelled in the macroeconomic model at an aggregate level, while in the credit risk model default risk is modelled at the borrower level.
The risk of model inconsistency is only likely to arise in the context of modular approaches since, in principle, integrated approaches are internally consistent in their construction. However, at least at present, the integrated modelling approach comes at the cost of reducing the model’s realism and intuitive interpretation. That is, models with integrated approaches tend to have financial sectors or credit relationships that are rather simplistic, therefore often lacking realism. In contrast, the modular approach allows adding more realism by introducing modular extensions that capture several types of bank risk (such as market, liquidity and sovereign risk in addition to credit risk) and important features of risk generation and propagation in the financial sector (such as interbank market contagion, asset fire sales and liquidity risk channels). Working with different modules may also facilitate robustness checks on different parameters and assumptions within each module, potentially providing a range of stress testing outcomes corresponding to different underlying parameters and assumptions. The risk of model inconsistency increases with the number of modules included in the model, however, especially as the different modules are often linked by reduced-form equations or rules of thumb.

Box 2 – Level of data aggregation

This Box considers the different levels of aggregation at which credit risk can be linked to the macroeconomic environment. The different levels of data aggregation are illustrated in the Chart below, which considers the corporate exposures of a set of J banks to firms in N industrial sectors in a particular jurisdiction (country). In our discussion, we focus on four possible levels of aggregation: the borrower (firm) level, the bank level, the sectoral level and the aggregate level.

The most granular approach in linking credit risk parameters to macroeconomic developments can be obtained using borrower-level data (e.g., estimating the borrower’s probability of default). In this approach, the credit risk variable for each separate firm (a, b, c, etc.) in the Chart is linked to macroeconomic variables. By allowing the sensitivity of credit risk variables to macroeconomic variables to differ across firms, differences in the potential reactions of firms to adverse macroeconomic shocks may be revealed. This would also allow a more detailed picture to be obtained of the effect of macroeconomic stress on a bank’s portfolio. However, in order to fully exploit the informational advantage obtained by using borrower-level data, the person conducting the stress test should have knowledge of (or make assumptions about) the exact composition of the banks’ portfolios. This may be problematic for top-down stress tests. Even in a bottom-up stress test the borrower-level approach may not be

---

(1) The risk of internal inconsistencies, both conceptual and empirical, arising from the modular structure of stress testing models has also been raised by Borio and Drehmann (2009), for example.

(2) Examples of such models include Boss et al. (2006), Aikman et al. (2009), Alessandri et al. (2009) and Gauthier et al. (2010).

(1) This need not coincide with the level of aggregation – e.g., portfolio, bank or system-wide level – at which the resilience assessment in terms of losses or capital adequacy is performed.

(2) As most of the papers in the credit stress testing literature focus on corporate credit risk, the illustration concentrates on firms and industrial sectors. However, the example may be generalised to other types of borrowers and portfolios.
feasible, in that evaluating the sensitivity of the credit risk of every individual borrower to macroeconomic stress may require substantial resources.

When data on the credit risk of individual borrowers and/or information on the banks’ portfolio composition are not available, an alternative may be to use credit risk information aggregated at the level of the different banks (bank-level data, such as non-performing loans or impairments), as indicated by the green dashed boxes in the Chart. While linking credit risk to the macroeconomic environment at this level of aggregation does not allow for differentiating between the effects of macroeconomic stress on the different individual firms or the different loan portfolios of the bank, it does allow for distinguishing the effects of macroeconomic stress across banks.

Alternatively, instead of vertical aggregation of the firm-level credit risk data in the Chart, the aggregation may also take place in a horizontal direction, to obtain sectoral credit risk data (see the blue dashed boxes in the Chart). The sectoral approach allows for a heterogeneous treatment of banks in terms of sensitivity to macroeconomic stress, provided that information is available on different banks’ exposures to different sectors. If this is the case, it may also be possible to obtain sectoral credit risk data at the bank level (bank-sector observations), i.e., horizontally aggregating firms for each bank separately as in the purple dashed boxes in the Chart.

The final level of aggregation is economy-wide, as indicated by the red dashed box in the Chart. The advantage of such an aggregate approach is that economy-wide data are generally more readily available than more granular data. Modelling the aggregate behaviour of borrowers is also potentially less complex than the individual borrower approach. Nevertheless, aggregate data potentially ignore significant variations across firms or banks that would be captured in a more granular approach.

In summary, an important question relating to data aggregation is whether stress testing outcomes are robust across different levels of aggregation. (1)

(1) Düllmann and Kick (2010) compare the results of a credit stress test based on borrower-specific PDs to those obtained for a sectoral average of the PDs. They find a substantial information gain in using borrower-specific PDs instead of sector-level PD, i.e., there is a higher dispersion across banks in the relative increase in expected losses under the stress scenario in the case of using borrower-specific PDs compared to sector-level PDs.

2. The choice of the credit risk variable

In the previous Section, we deliberately remained general in our discussion of credit risk by referring to credit risk parameters or credit risk variables. In this section, we discuss the specific variables that could be used to measure credit risk. In general, three concepts play a crucial role in credit risk stress testing: the probability of default (PD), the loss given default (LGD) and the exposure at default (EAD). This terminology, which reflects the Basel II framework, allows computing, on a portfolio basis, expected loss as $EL = PD \times LGD \times EAD$. The expected loss estimate is based on a best estimate of the deterioration in the creditworthiness of the credit portfolio or of default (i.e., an increase in PD and/or LGD). (1)

In order to implement a stress test, the credit risk components of expected losses have to be mapped into observable variables. The basic question that arises in this context is which variable(s) to use for measuring credit risk in the portfolio. In the remainder of this Section we discuss the different types of credit risk variables used in practice and raise the issue of robustness of stress testing outcomes to the choice of different credit risk variables. This choice is often determined by practical considerations, such as data availability at the desired level of aggregation (see Box 2 for an overview of the different levels of aggregation at which credit risk can be linked to the macroeconomic environment).

2.1 Credit risk variables used in practice

Several variables – some of which are purely aimed at capturing PDs, while others may also include LGD components (2) – have been used in credit stress testing

(1) EAD is regularly reported by banks and is typically treated as fixed over the stress testing horizon.

(2) Whereas in stress testing applications, the stressed values of PD are usually obtained from an estimated equation that explains PD using macroeconomic variables as explanatory variables, the stressed value of LGD is often based on an assumption (e.g., imposing an ad-hoc value for the LGD over the stress testing horizon).
applications. We distinguish between variables based on bank accounting data, default data and model-based measures, respectively. The main characteristics of these variables are also summarized in Table 2.

2.1.1 Bank accounting data

A bank accounting measure that may serve as a proxy for default rates is the ratio of (new) non-performing loans to a given measure of total loans (NPL ratio). Conceptually, the NPL ratio may be a good proxy for PD, as non-performing loans are those loans for which it is likely that a contractual payment will not be made and which are therefore technically in default. However, there is no single definition of non-performing loans; while non-performing loans are usually determined by a criterion such as x days overdue, the value of x may differ across banks (typically x equals 60 or 90). Therefore, unless some harmonisation has been imposed by regulatory authorities (as in Belgium) the NPL ratio may not be entirely comparable across different banks.

Following an increase in expected credit loss (due to a rise in PD or LGD), banks may create new provisions and record them as impairments on the expense side of their P&L with an ultimate impact on their Tier 1 capital through lower retained earnings. Therefore, another natural candidate for capturing credit risk is the ratio of (new) loan loss provisions (or impairments) to a given measure of the stock of total loans (LLP ratio, also referred

(1) This is suggested for example by Blaschke et al. (2001).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Content</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bank accounting data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-performing loan (NPL) ratio</td>
<td>Ratio of NPLs to a measure of total loans; NPLs are loans for which principal or interest payments are more than x days in arrears, where x typically equals 60 or 90</td>
<td>PD</td>
<td>Potential absence of harmonized definition for NPLs across banks; affected by write-offs</td>
</tr>
<tr>
<td>Loan loss provision (LLP) ratio</td>
<td>Ratio of LLPs to a measure of total loans; new provisions may be created and recorded as impairments, following an increase in expected losses on a loan portfolio (potentially before arrears in payments have occurred)</td>
<td>PD; LGD</td>
<td>Dependent on accounting rules and bank’s discretion/provisioning policy; affected by write-offs and reversals</td>
</tr>
<tr>
<td><strong>Default data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default rate</td>
<td>Ratio of number of borrowers in default to total number of borrowers; under the Basel II framework, an obligor is considered to be in default when either the bank considers that the borrower is unlikely to pay its credit obligations in full or when any material credit obligation to the bank is more than 90 days overdue</td>
<td>PD</td>
<td>Not publicly available</td>
</tr>
<tr>
<td>Bankruptcy rate</td>
<td>Ratio of the number of firms filing for bankruptcy to total number of firms</td>
<td>PD</td>
<td>Understates default; generally applies only to firms</td>
</tr>
<tr>
<td><strong>Model-based measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market-based measures</td>
<td>Probability (or expectation) of default over a future one-year horizon extracted from market information (e.g., stock prices, CDS spreads)</td>
<td>PD</td>
<td>Only available for publicly listed firms; market prices are affected by non-credit risk related factors</td>
</tr>
<tr>
<td>Banks’ internally estimated PDs</td>
<td>Probability (or expectation) of default over a future one-year horizon based on bank’s internal model</td>
<td>PD</td>
<td>Not publicly available; limited history; bank model dependent; potentially a “through-the-cycle” measure</td>
</tr>
</tbody>
</table>
to as the credit cost ratio). Interestingly, the LLP ratio is not a pure measure of PD; it also entails an LGD component. However, accounting rules on provisioning may differ across jurisdictions. In addition, changes in the LLP ratio do not necessarily reflect only credit risk, but may also depend on banks’ provisioning policies, which may involve, for example, income smoothing for fiscal reasons. Thus, within the existing regulations, banks often have some discretion in recording new impairments.

This potential lack of comparability of non-performing loans and loan loss provisions across banks implies that it may be useful to complement bank accounting measures of credit risk with more direct or strictly defined measures of borrower default to assess their robustness.

2.1.2 Default data

A direct measure of PD may be found in banks’ observed default data; these data may be available at borrower level (as a 0/1 indicator of default), or at more aggregate levels expressed as a default rate. In contrast to non-performing loans, the definition of default is harmonised under the Basel II framework, according to which an obligor is considered to be in default when either the bank considers that the borrower is unlikely to pay its credit obligations in full or when he is more than 90 days overdue on any material credit obligation to the bank. Banks’ default data are sometimes also centralized in (central banks’) credit registers, to which banks typically have to report loan-level information for each loan contract whose size exceeds a given (relatively low) threshold. These data sources therefore provide wide coverage of the banks’ loan portfolios. In contrast to bank accounting data, however, these data are usually not publicly available.

As an alternative to default information, which is not always available in credit registers, information on bankruptcy filings may be used. While bankruptcy rates are likely to underestimate default – as default is not always followed by bankruptcy – they are nevertheless relatively frequently used as a credit risk measure in stress testing because of the lack of publicly available data on default. In contrast to default data, information on bankruptcy filings is usually publicly available. However, bankruptcy data usually only apply to firms; therefore, they cannot provide an estimate of default rates for loans to households, for example.

2.1.3 Model-based measures

A final category of measures used as credit risk variables in stress testing consists of model-based measures. Rather than directly conducting the analysis on the basis of observed credit risk variables, these studies typically employ a combination of market information, such as stock returns or CDS spreads, and a (structural) portfolio credit risk model to generate estimates of banks’ credit losses under adverse scenarios.

When market information is used to derive the credit risk variable, two possible approaches can be followed to link the credit risk variable to the macroeconomic variables. One approach is to first estimate the credit risk component of the market price, using macroeconomic variables as explanatory variables, and then to evaluate the credit risk component using stressed values of the macroeconomic variables. An alternative approach is to first estimate the relationship between the market data and the macroeconomic variables, to evaluate the market prices at stressed values of the macroeconomic variables, and then to extract the credit risk variable from the stressed market price data. An interesting but unanswered question is how big the difference is between the stressed values of the credit risk variables, and hence, the outcomes of the stress test, from these two approaches.

The advantage of the market-based indicators is that they can be constructed on the basis of borrower-level data that are publicly available. However, the coverage of banks’ loan portfolios that can be obtained in this approach may be limited, since market data are only available for publicly listed firms and are typically not available for SMEs and households, which may represent a large fraction of the banks’ loan portfolios. Finally, movements in market prices are not necessarily related to credit risk; market price changes are also likely to reflect other factors that may be unrelated to credit risk, such as market liquidity, investors’ risk aversion or general market sentiment.

As an alternative to market-based credit risk information, one may consider using banks’ model-based internally estimated PDs, which are assigned to each loan in a portfolio. In the internal ratings-based (IRB) approach of the Basel framework, these estimated default probabilities are used for computing risk-weighted assets in order to determine the bank’s regulatory capital requirement. As banks’ internally estimated PDs are usually not publicly available and have a short history, these data are rarely used in stress testing exercises. In addition, some authorities have found that different banks sometimes assign quite different PDs to the same firms. Another disadvantage of

---

(1) In fact, van den End, Hoeberichts and Tabbaa (2006) exploit this feature of the LLP ratio to separately identify the behaviour of LGD in addition to that of PD by combining data on the LLP ratio with bankruptcy rate data.

(2) See e.g., Foglia (2009).

(3) This is not only true for stock prices, which in general may be expected to depend on all factors that affect the firm’s future profitability, but also for credit default swaps (see e.g., Collin-Dufresne et al. (2001), Annaert et al. (2010) and Bongaerts et al. (2011)).
banks’ internally estimated default probabilities, besides their limited availability, is the fact that these PDs are often “through-the-cycle”, meaning that the PD for a borrower may be estimated either as an average value over an entire cycle or as the probability of default in a downturn.

2.2 Robustness of stress testing outcomes across different credit risk variables

As no ideal measure of credit risk seems to exist, practitioners often have to choose among different potential credit risk variables, and the ultimate choice may to a large extent be driven by practical considerations such as data availability. It is therefore important to assess the robustness of stress testing results to the choice of the credit risk variable. In this sub-Section we discuss some of the factors that may contribute to differences in stress testing outcomes when using different credit risk variables. In particular, we focus on the issue of using stocks versus flows and on the speed at which the different credit risk variables incorporate credit risk information. In Box 4 we provide an illustration of the issue of robustness of stress testing results across different credit risk variables using Belgian data.

2.2.1 Stocks versus flows

When using bank accounting information (non-performing loans, loan loss provisions) to construct a credit risk variable, it has been argued that write-offs (or charge-offs) and reversals (or recoveries) may blur the picture of developments in credit risk obtained from the accounting data. (1) For example, large write-offs may result in a decrease in the stock of LLPs (and NPLs), even when the flow of new provisions (and NPLs) has increased. Some practitioners have therefore suggested using the flow of new NPLs or LLPs rather than the stock of existing NPLs or LLPs, or adjusting the stocks for write-offs. (2)

In addition, the composition of the loan portfolio may change over time, as existing loan contracts end and new contracts enter the banking book. As NPLs and LLPs are typically expressed relative to some measure of total loans (the NPL ratio and the LLP ratio), strong loan growth may result in a decrease in the LLP (and NPL) ratio, even when the flow of new provisions (and NPLs) may have increased. Similar comments may also apply to alternative credit risk variables, such as default rates and bankruptcy rates.

In summary, the use of different variants of particular credit risk variables may introduce sources of divergence in stress testing outcomes, as for example an LLP ratio constructed from the stock of LLPs may react differently to the macroeconomic variables than an LLP ratio based on the flow of new provisions.

2.2.2 Speed of credit risk information incorporation

Another feature of the data that may affect the robustness of stress testing outcomes to the choice of credit risk variable is the speed at which credit information is reflected in the credit risk variable. One issue that has received some attention in stress testing is the backward or forward-looking nature of the data. On the one hand, event-based data, such as non-performing loans, and default and bankruptcy data, are typically backward-looking, in that they include information only on past events. Moreover, the speed at which the evolution of credit risk is reflected in these event-based variables depends on the underlying definitions of the event, which may differ across the different variables (e.g., the number of days overdue for non-performing loans and default data, the time between actual distress and the bankruptcy filing in the case of bankruptcy rates). Credit risk variables based on loan loss provisions may, in principle, also be considered to be backward-looking, but the extent to which this is the case may depend on the bank’s provisioning policy.

Model-based credit risk variables, such as those based on market information or banks’ internally estimated PDs, are said to be forward-looking, or ex ante, measures of credit risk. Model-based PDs (if based on the Basel II concept) typically reflect the probability (or expectation) of default over a future one-year horizon. Important in this respect is also the potential through-the-cycle nature of the model-based credit risk variables. (3) To the extent that they are indeed through-the-cycle, as opposed to “point-in-time”, this may hamper the finding of any statistically significant relationship of the credit risk variable with macroeconomic movements. (4)

These timing features of the various credit risk variables may imply different (timing in the) reactions of specific credit risk variables to stressed macroeconomic variables and hence, potentially different stress testing outcomes.

---

(1) Write-offs imply the removal of a loan from the balance sheet when the loan is considered completely unrecoverable. Reversals occur on impairments that have previously been recognised in the accounts and follow events that suggest an increase in the recoverable amount of the asset (e.g., an improvement of the borrower’s ability to pay). Reversals are recognised as a credit to the P&L.

(2) See e.g., Pain (2003) and Jakubík and Schmieder (2008).

(3) See Section 2.1.3.

(4) Sorge and Virtanen (2006) observe that while accounting measures of risk are very sensitive to the business cycle, market-based indicators exhibit substantial variability both across firms and over time, but appear to be less responsive to macroeconomic or systematic risk factors.
Box 3 – Illustration of robustness issue using data for Belgian banks

This Box assesses the impact of the choice of the credit risk variable on the outcome of a credit stress testing exercise. We consider the impact of a hypothetical macroeconomic scenario on three potential credit risk measures for the Belgian banking sector: the non-performing loans ratio (NPL ratio), the flow of new impairments over total loans (LLP ratio), and the bankruptcy rate. (1)

The results below are based on highly simplified modelling assumptions, purely to illustrate the potential impact of the choice of the credit risk variable. These results should not be interpreted as the outcome of a real credit stress testing exercise for the Belgian banking sector. In particular, we relate quarterly observations of the credit risk variables to their own lagged values and to current and lagged values of the yearly GDP growth rate of Belgium and the 10-year yield of Belgian government bonds. To facilitate comparison of the variables and results, we estimate the same specification for each of the three credit risk variables. (2) For estimation of each of the models, we consider a common sample for the three data series spanning the time period 1995Q1-2009Q4.

After estimation of the model, we assume that for the period 2010Q1-2011Q4 history repeats itself and that GDP growth rates and long-term interest rates follow the same path as they did over the crisis period 2008Q1-2009Q4. This historical scenario is selected on an ad-hoc basis and is aimed only at providing an illustration of the (potential lack of) robustness of stress testing outcomes for different credit risk variables. The left-hand panel of the Chart below shows the observed and stressed pattern (projection 2010-2011) for the macroeconomic variables (in %), (left-hand scale)

(1) The NPL and LLP ratios were obtained from data in the supervisory reporting scheme. Bankruptcy rates were obtained from data on bankruptcy filings and the NBB’s central corporate credit register.

(2) The credit risk variables were seasonally adjusted, logit transformed and first-differenced before estimating the credit risk equation, in which four lags were included for each explanatory variable.
whereas the right-hand panel presents the observed (normalised) credit risk variables and the calculated impact of the historical macroeconomic scenario for the three credit risk variables (projection 2010-2011)(1).

From the right-hand panel, we observe that the reactions of the three credit risk variables to the macroeconomic stress scenario seem to differ somewhat. While both the bankruptcy rate and the NPL ratio increase from the first quarter onward, the LLP ratio drops during the first quarter of the stress horizon and then increases again. Moreover, while bankruptcy rates and the LLP ratio attain their maximum value in the third quarter of 2011, the NPL ratio reaches its maximum value earlier, in the fourth quarter of 2010. We also see that bankruptcy rates seem to react somewhat more strongly to the imposed stress scenario. More precisely, the maximal change over the period 2010Q1-2011Q4 is 18.9 % and 36.1 %, respectively for impairments and for bankruptcy rates, and 10.5 % for NPL. One of the potential reasons for these differences in the patterns of stressed values for the credit risk variables may be found in variations in the timing with which variables reflects changes in credit risk. Another explanation may be the different coverage of the variables: whereas bankruptcy rates only include exposures to Belgian firms, the LLP and NPL ratio include exposures to both firms and households, as well as foreign exposures of Belgian banks.

Finally, we provide a back-of-the-envelope calculation to illustrate the potential implications of the above differences for stressed values in terms of P&L and bank capital. Assuming that LGD is fixed at a value of 0.45, we can interpret the reported stressed increases in the credit risk variables as proxies for stressed increases in PD. Given that $EL = PD \times LGD \times EAD$ and that the major Belgian banks have an average EAD to the corporate and retail sector of € 152bn, our hypothetical macroeconomic stress scenario would increase EL by € 165mln based on the NPL ratio, € 402mln based on bankruptcy rates and € 31mln based on the LLP ratio. These increases in EL amount to about 1.02 %, 2.50% and 0.19 %, respectively, of average Tier 1 capital held by the major Belgian banks.

In summary, the choice of the credit risk variable may matter for the outcome of a stress test. The extent to which this choice results in large potential differences in terms of P&L and balance sheet effects should be investigated on a case-by-case basis.

(1) A deterministic approach is used to obtain the stressed values of the credit risk variables. To facilitate comparison, the credit risk variables are normalised such that their value in 2009Q4 equals 100.

3. Concluding remarks

In this article we have discussed a number of questions related to model risk in credit stress testing that have not been widely discussed but that nevertheless need to be addressed and understood by practitioners of stress tests. We have focused on the relative implications of differing modelling choices in the stress testing process and, in particular, on the questions of how to link credit risk to the macroeconomic environment and which variables to use for measuring credit risk in the portfolio.

Regarding the link between the macroeconomic environment and credit risk, we distinguish between integrated and modular approaches. An integrated approach allows consistent treatment of all relationships between macroeconomic variables and credit risk variables, but at the cost of reducing the model’s realism and intuitive interpretation. In contrast, the modular approach allows adding more realism and may facilitate robustness checks on different parameters and assumptions within each module, potentially providing a range of stress testing outcomes corresponding to different underlying parameters and assumptions. However, modular approaches embody the risk of model inconsistency.

With respect to the choice of credit risk variables, we argue that, since the ultimate choice may to a large extent be driven by practical considerations such as data availability, it is important to assess the robustness of stress testing results to the choice of the credit risk variable. In this context, we identify potential factors that may contribute to variations in stress testing outcomes when using different credit risk variables. In addition to the alternative definitions of various proxies for default, we focus on the question of using stock versus flow variables and on the issue of the varying speeds at which the different credit risk variables incorporate credit risk information.
References


Düllmann, K. and T. Kick (2010), Stress Testing German Banks Against a Global Credit Crunch, Deutsche Bundesbank, mimeo.


Głogowksi, A. (2008), Macroeconomic Determinants of Polish banks’ Loan Losses – Results of a Panel Data Study, National B
## Overview of and references to the different approaches

### Modeling framework

<table>
<thead>
<tr>
<th>Approach</th>
<th>References</th>
</tr>
</thead>
</table>

### Level of data aggregation

<table>
<thead>
<tr>
<th>Level</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borrower-level</td>
<td>Drehman (2005), Pesaran et al. (2006), Andersen et al. (2008), Bernhardsen and Syversten (2009), Düllmann and Kick (2010)</td>
</tr>
</tbody>
</table>

### Choice of credit risk variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default rate</td>
<td>Wong, Choi and Fong (2008), Jiménez and Mencia (2009), Simons and Rolwes (2009)</td>
</tr>
<tr>
<td>Banks’ internally estimated PDs</td>
<td>Düllmann and Kick (2010)</td>
</tr>
</tbody>
</table>