PD estimates for Basel II

One of the main issues banks will have to face to comply with the new Basel II internal ratings-based approach is to prove that the long-run average probabilities of default they assign to their clients, which will be used as the basis for regulatory capital requirements, are correct. Currently, there are no standard tests to compare them with observed default rates. Laurent Balthazar develops an approach that is directly derived from the Basel II theoretical framework and proposes tests that could be a basis of discussion between banks and regulators

Basel II is completely changing the way banks will have to calculate their regulatory capital requirements. The new regulations are much more risk-sensitive than the previous ones, as capital will become a function of (among other things) the risk that a counterparty does not meet its financial obligations. In the standard approach, risk is evaluated through external ratings given by recognised rating agencies. In the internal ratings-based (IRB) approach, banks will have to estimate a probability of default (PD) for each of their clients. Of course, to qualify for IRB, banks will have to demonstrate that the PDs they use to calculate their risk-weighted assets are correct. One of the tests required by regulators will be to compare the estimated PDs with observed default rates (DR). This will be difficult as DRs are usually very low and highly volatile.

The aim of this article is to show how we can develop hypothesis tests that could help to make this comparison. Credit risk models and default generating processes are still subject to much debate in the industry, as there is no consensus on what is the best approach. So, we have chosen to build the tests starting only from the simplified Basel II framework. The reason is that, despite the criticism it has met, it will be mandatory for major US banks and almost all European ones. So, even if many find it too simplistic, banks will have to use parameters (PDs and others) that give results consistent with observed data, even if they think that bias arises from model mis-specifications. Our goal is to propose simple tools that could be some of the many used as a basis for discussion between banks and regulators during the validation process.

PD estimates

Banks will be required to estimate a one-year PD for each obligor. Of course, 'true' PDs can be assumed to follow a continuous process and might thus be different for each counterparty. But in practice, true PDs are unknown and can only be estimated through rating systems. Rating systems can be statistical models or expert-based approaches (most of the time they are a combination of both) that classify obligors in different rating categories. The number of categories vary but tend to lie between five and 20 (Basel Committee, 2000). Companies in each rating category are then supposed to have relatively homogeneous PDs (at least the bank is not able to discriminate further). Estimated PDs can sometimes be inferred from equity prices or bond spreads, but historical default experience will usually be used as the most reasonable estimate. So, banks will have groups of counterparties that are in the same rating class (and then have the same estimated PD derived from historical data) and will have to check if the DRs they observe each year are consistent with their estimation of the long-run average one-year PD.

Basel II framework

The new Basel II capital requirements have been established using a simplified portfolio credit risk model. The philosophy is similar to market standards KMV and CreditMetrics but in a less sophisticated form. In fact, it is based on the Vasicek one-factor model (1987), which builds upon Merton's value of the firm framework (1974). In this approach, asset returns of a company are supposed to follow a normal distribution. We will not discuss the presentation of the model further, as it has already been extensively documented (see, for instance, Finger, 2001).

If we do not consider all the formula but only look at the part related to PD, the Basel formula is the following: for a PD and an asset correlation ρ, the required capital is:

$$\text{Regulatory capital}^1 = \Phi\left(\frac{\Phi^{-1}(PD)+\sqrt{\rho}\times\Phi^{-1}(0.999)}{\sqrt{1-\rho}}\right) - PD \qquad (1)$$

Φ and Φ^{-1} stand respectively for the normal and inverse normal standard cumulative distributions. The formula is calibrated to calculate the maximum default rate at the 99.9th percentile (and the average PD is subtracted to focus on unexpected loss only). With elementary transformations, we can construct a confidence interval (CI) at the α level (note that the formula for capital at the 99.9th percentile is based on a one-tailed test while the constructed CI is based on a two-tailed test):

$$\left[\Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\times\Phi^{-1}(1-\alpha/2)}{\sqrt{1-\rho}}\right);\Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\times\Phi^{-1}(\alpha/2)}{\sqrt{1-\rho}}\right)\right] \qquad (2)$$

So, the formula above gives us a CI for a given level of PD if we rely on the Basel II framework. For instance, if we expect a 0.15% DR on one rating class, using the implied asset correlation in the Basel II formula (23.13% if we assume we test a portfolio of corporates, see Consultative Paper 3 for details), and a CI at 99% level ($\alpha = 1\%$), we get the following: [0.00%; 2.43%]. It means that if we observe a DR that is beyond those values we can conclude that there is a 99% chance that there is a problem with the estimated PD.

As already discussed above, one could argue that the problem does not come from the estimated PD but has other causes: wrong asset correlation level, wrong asset correlation structure (the one-factor model should be replaced by a multi-factor model), rating class is not homogeneous in terms of PDs, bias from small sample size (we will discuss how to deal with this below), wrong assumption of normality of asset returns, etc. Anyway, the Basel II model will be mandatory. So, a bias due to too-weak correlation implied by the Basel II formula, for instance, should be compensated by higher estimated PDs (or by additional capital required by regulators under the second pillar of the Accord).

Correction for finite sample size

One of the problems banks and regulators will have to face is the small sample of counterparties that constitute some rating classes. The Basel II formula is constructed to estimate stress PDs on infinitely granular portfolios (where the number of observations tends to infinity). If an estimated PD of 0.15% applies only to a group of 150 counterparties, we can imag-

[1] *Without taking into account maturity adjustment (formula is for the one-year horizon) and for LGD and EAD equal to 100%*

ine that the variance of observed DR could be higher than the one forecast by the model.

Fortunately, this bias can easily be incorporated in our construction of a CI by using Monte Carlo simulations. Implementing the Basel II framework can be done through a well-known algorithm:

- Generate a random variable $X \sim N(0, 1)$. It represents a common factor to all asset returns.
- Generate a vector of the n Y random variable (n being the number of observations a bank has on its historical data) with $Y \sim N(0, 1)$. It represents the idiosyncratic part of the asset returns.
- Calculate the firms' standardised returns as:

$$\begin{bmatrix} Z_1 \\ \cdots \\ Z_n \end{bmatrix} = \sqrt{\rho} X + \sqrt{1-\rho} \begin{bmatrix} Y_1 \\ \cdots \\ Y_n \end{bmatrix}$$

- Define the returns thresholds that lead to default as $T = \Phi^{-1}(PD)$.
- Calculate the number of defaults in the sample as:

$$\sum D_i \begin{cases} D_i = 1 \text{ if } Z_i < T \\ D_i = 0 \text{ if } Z_i \geq T \end{cases}$$

- We can then calculate the average DR in the simulated sample.
- We repeat the above steps, say, 100,000 times and we will get a distribution of simulated DR with the correlation we assumed and incorporating the variability due to our sample size. Then, we only have to select the desired α level.

Extending the framework

We still face an important problem: the CIs are too wide. For instance, for a 50-basis-point PD, a CI at 99% would be [0.00%; 5.91%]. So, if a bank estimates 50bp of PD on one rating class and observes a DR of 5% the following year, it still cannot reject the hypothesis that its estimated PD is too weak. But we can intuitively understand that if, over the next five years, the bank observes 5% of DR each year, its 50bp initial estimates should certainly be reviewed. So, if conclusions about the correct evaluation of the PD associated with a rating class are hard to check with one year of data, several years of history should allow us to draw quicker conclusions (Basel II requires that banks have at least three years' data before qualifying for the IRB).

In the simplest case, one could suppose that the realisations of the systematic factor are independent from one year to another. Then, extending the Monte Carlo framework to simulate cumulative DR is easy, as we only have after the seventh step to go back to step one and make an additional simulation for companies that are in the same rating class the following year. We do this t times for a cohort of t years and we can then calculate the cumulative default rate of the simulated cohort. We follow this process several thousand times so that we can generate a whole distribution and we can then calculate our CI.

As an example, we have run the tests with the following parameters: number of companies = 300, PD = 1%, correlation = 19.3% and number of years = five. This gives us the following results for the 99% CI: one year – [0.0%; 9.7%]; two years – [0.0%; 12.7%]; three years – [0.0%; 15.7%]; four years – [0.0%; 17.7%]; and five years – [0.3%; 19.3%].

To see clearly how the effect of the cohort approach allows us to narrow our CI, we have transformed those cumulative CI in yearly CI using the Basel II proposed formula:

$$PD_{1\,year} = 1 - \left(1 - PD_{t-years}\right)^{(1/t)}$$

where one year – [0.0%; 9.7%]; two years – [0.0%; 6.5%]; three years – [0.0%; 5.4%]; four years – [0.0%; 4.6%]; and five years – [0.06%; 4.2%].

We can see that the upper bound of the annual CI decreases from 9.7% for one year of data to 4.2% for five years of data. It shows that the precision of our hypothesis tests can be significantly improved when we have several years of data.

Of course, those estimates could be undervalued because one could reasonably suppose that the realisations of systematic factors are correlated from one year to another. This would result in wider CIs. It could be an area of further research that is beyond the scope of this article, as modifying the framework to take account of this could be done in many ways and cannot be directly constructed from the Basel II formula.

But we can conclude that regulators should not allow a bank to lower its estimated PD if the lower bound is broken, while they could require a higher estimated PD on the rating class if observed DR is above the upper limit.

Conclusions

In the Federal Reserve's (2003) draft paper on supervisory guidance for the IRB, we find the following: "Banks must establish internal tolerance limits for differences between expected and actual outcomes... At this time, there is no generally agreed-upon statistical test of the accuracy of IRB systems. Banks must develop statistical tests to back-test their IRB rating systems..."

In this article, we have tried to answer the following question: which level of observed default rate on one rating class should lead us to have doubts about the estimated probability of default we use to calculate our regulatory capital requirements in a Basel II context?

Many approaches can be used to describe the default process, but we have decided to focus on the Basel II proposed framework. Though it is currently subject to debate, the final framework will be imposed on banks. Then the parameters used should deliver results consistent with the regulators' model.

We have shown how to construct a hypothesis test using a confidence interval derived directly from the formula in the Basel Consultative Paper 3. We have explained how we can build a simple simulation model that gives us results that integrate variance due to the size of the sample (while the original formula is for the infinitely granular case). Finally, we have explained how to extend the simulation framework to generate a cumulative default rate under the simplifying assumption of independence of systematic risk from one year to another. This last step is necessary if we want to have a CI of a reasonable magnitude.

This approach could be one of the many used by banks and regulators to discuss the quality of the estimated probabilities of default. The model has been implemented in a VBA program and is available from the author upon request. ■

Laurent Balthazar works in the credit risk methodology department of Dexia Bank in Belgium. The views expressed here are those of the author and do not necessarily reflect the opinion of Dexia Bank. Email: Laurent.Balthazar@Dexia.be

REFERENCES

Basel Committee working group, 2000
Range of practice in banks' internal rating systems
Available at *www.bis.org/publ/bcbs66.pdf*

Federal Reserve, 2003
Internal ratings-based systems for corporate credit and operational risk advanced measurement approaches for regulatory capital
Available at
www.federalreserve.gov/boarddocs/press/bcreg/2003/20030804/attachment2.pdf

Finger C, 2001
The one factor model in the new Basle Capital Accord
Available at *www.riskmetrics.com/pdf/journals/rmj2.pdf*

Merton R, 1974
On the pricing of corporate debt
Journal of Finance 29, pages 449–470

Vasicek O, 1987
Probability of loss on loan portfolio
Available at *www.moodyskmv.com/research/whitepaper/Probability_of_Loss_on_Loan_Portfolio.pdf*